

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAAA
 GGAAAGTGTGGAGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACCTC
 AGTGAATGAGGCCAGGCAGGAAACGCGCGGAAGGATTGTGAGCTTGGGACCCCT
 TCATAGACACTGATGACACGTTACGCAAAATAGAAATTGAGGAGAAACGCCCTGG
 GCCTCGGAAAGGAGTGATTGATTAGTACTTGAAGTTAGGTGACTTTAAGGAGAA
 CTAACATAATGTATACTATTGAGGGAGGAGGAAGGAGCATTACAGAGTTCCAGCAGC
 AGCAGGAAAGCTTGGTTAATTGAAATGGATGATAGCATTAAAATAACAGAAGC
 GCCTCCAGGTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
 CACTAAGCCTTTGATCCCTTGGAAAGCAAAGAACTTCCCTCCCTGGGTGAAGAC
 TCTCCTCAGAAGATTCCCTGCTCTGCCTATGTTACAAGAGGAATCAAACCAAGAC
 AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAAGTCAGCGGTTGATCTCCC
 TCACTGCTGCTGCCTGGTGCACCCCTGGGAGCAGGGCCTGCTCGCCGCTGTG
 CCTGCTATGTGCCACAGAGGTGACTGTACATTGGTACCTGACCTCCATCCCAG
 ATGGCATCCCAGGAAATGTGGAACGAATAAATTAGGATATAACAGCCTTACTAGAT
 TGACAGAAAACGACTTGATGGCCTGAGCAAACCTGGAGTTACTCATGCTGACAGT
 AATGGCATTACAGAGTCAGTGACAAGACCTCTGGGCTTGAGTCCTGCAGGTC
 TTAAAAATGAGCTATAACAAAGTCAAATCATTGGAAAGGATACTTCTACGGACTC
 GGGAGCTGGTCCGGTGCACCTGGATCACAACAAATTGAATTCATCAACCTGAG
 GCCTTATGGACTACCTCGCTCCGCTGGTACATTAGAAGGAAACCGGCTCACA
 AAGCTCCATCCAGACACATTGTCTCATTAAGCTATCTCCAGATATTAAAACCTTT
 TCATTAAGTACCTGTTCTGCTGATAACTTCCCTGACCTCCCTCCAAAAGAAATGGT
 CTCCTACATGCCAAACCTAGAAAGCCTGTATTGATGGAAACCCATGGACCTGTGA
 CTGCCATTAAAGTGGTGTGAGTGGATGCAGGGAAACCCAGATATAATAAAAAT
 GCAAGAAAGACAGAAGCTTCCAGTCAGCAATGTCCCCTTGATGAACCCCA
 GGATCTCTAAAGGCAGACCCCTTGCTATGGTACCATCTGGAGCTTCTATGTACAA
 AGCCAACCATGATCCATCACTGAAGTCAAAGAGCCTGGTACTCAGGAGGACAAT
 GGATCTGCCTCCACCTCACCTCAAGATTGATAGAACCCCTTGGCTCTGTCTTGA
 ACATGACANANNTNTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
 TCAAGGACATCACCACACTGCAATTGAGAAGAAAATGACTACATCATGCTAAATGC
 GTCATTTCACAAATCTGTGTCAGTGTAGATTATAATCACATCCAGCCAGTGTG
 GCAACTTCTGGCTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
 TACCGAGACTCCTTCACTGTCTCTAGATATAAACAGGTGGCTCTAGGCTGAAGA
 CATTTTACCAAGCATAGAGGCTGATGTCAGAGCAGACCCCTTGGTCCAAACAAGA
 AAAAATTGTCTGAGCTGAACAGAACTGCCACACACTAGCACATTACAGATCCA
 GTTTCCACTGATGCTCAAATCGCTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
 TCAAATGGACCATGATCCTGATGATGAAACAATCCAAACTGGAACGCACGTGCTGG
 TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCCTCACCTCACTGG
 AATGGCTCTAGCTGATGGGAGTAAAGTGAAGAGCCCTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAATGGGAAGTGGAACTGCAGATGGCTGACAGCTTGA
 TGCAGGTCTTACCACTGCATAAGCACCACATGATGCAGATGCGGATGTTCTCACATA
 CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
 ACACAGTGGTACGGGTGAGACGCTCGACCTTCATGCCTTCCACGGGTGTTCCAG
 ATGCTTCTATTAGCTGGATTCTCCAGGGAACACTGTGTTCTCAGCCATCAAGAG
 ACAGGCAAATTCTAACATGGACCTAACAGAATATTACAGGTACGCCAAAGAT
 CAAGGTCAATTACCAATGTGTGGCTGCCAACCCATCAGGGGCCACTTTCCAGTTT
 AAAGTTCAAGTCAAAAGAAAGGCCAAGGATGGTGAGCATGACAGGGAGGCAG
 GTGGATCTGGACTTGGAGAACCCAACCTCCAGTGTTCCTAACAGCAGCCAGCATCTT
 TGAAAACCTCTGCATCAGCTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGT
 TACATAGGAAGAACAAACATAGAGACTAACATACATCGGCGCGTGGGATTCCACG
 CTCCGGCGATTCAAGGGAGCATAGGAGGCAGCTCCCTCTCTGCTCGGAGAATTGAC
 CCGCAACGCTGGCAGCACTCTAGAAAAAGCCAAAAGAATTCTGTGCCAAAAAA
 GCAAGAAAATACCACAGTAAAGCCAGTGCCTGGCATGATTCTCCAGATGAAGAAATT
 CTGGTTC
 TGAAAACTAAGGCTCTGGTGTCCCAGGAAGGTACCAACTGCTGACTCTGGACCAG
 TAAATCATGGTTTATGACGAGTATAGCTCTGGCACAGAAGTCTCAACTGTGAAATC
 CACAAACACTACAATCTGAGCACCTCCTGATTCAAATTATTAGTGTAAACAAACG
 GTACAGCTGTGACAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATA
 ACCAACCAAAACCAATCATTATCTTCCATCAGTAGCTGAAATTGAGATTCTGCT
 CAGGCAGGAAGAGCATCTCCAAAGTGCACACCCCTGTAACAGGGGGAAACATGGC
 TACCTATGGCCATACCAACACATATAGTAGCTTACCAAGCCAGTACAGTCTT
 GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
 GCAGACCTAGCAGTAGTGACATCTCTCACACTACTGCAGACCCCTAGCTTCTCCA
 GTCACCCCTCAGGTTCACACACCACTGCCTCGTCTTATTACATTCTAGAAACAA
 CAATACAGGTAACCTCCCTGTCCAGGCACCTGGGAAGAGAGAGGACAATTGGA
 GCAGAGGGAGAGTTAAAACCCACATAGAACCCAGTTCTCCGACGGCATAGACAC
 AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTCC
 AGCCACAGAGTACCCCTGGGATGTGCCACACATGTCCCTCCGAGAGGGCTCACAG
 TGGCTACTGCAGCACTGTCAGTTCAAGTTCATCCCACAGTGCCTCCCCAAACTA
 ATAATGTTGGGTCAAGCAGAACAGACTTACCAACTGTGGTCAAGAAACCAACTGTAC
 TATTAAGGACAAACAAATGTAGATATTGAGATAATAACAACCAACTACAAATAT
 TCCGGAGGGAAAGTAACCACGTGATTCTACGGAAGCAAGCATGACTCTGCTCC
 AACATCTGTATCCCTGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
 TGGGACCATCCAAACTGGGAAAGATTCAAGTGGAAACAACACCACCTCCAGCCCC
 TCAGCACACCCCTCAATACCAACAAGCACAACATTCTCAAAGAGGAAACTCCCTG
 CACCAGATCTTGTAAATAACCAGAAGAAGAGGGATGTTAAAGAAATCCATATCA
 ATTGGTTACAAAAGAACCCAGCCGCAAAGCTCCAAAATAGCTCCTCTTITACC
 CACAGGTCAAGAGTTCCCCCTCAGATTCTACAACTCTTGCACAAGTCCGCCACCAGC
 TCTGTCTACAACAATGGCTGCCACTCAGAACAGGGACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAAGAAGCAGCCCCCACCAACTCCTCTCCAGTGCTTC
 CTAGCACCATAAGCAAGAGATCTAATACATTAAACTCTTGTCAACGGAAACCCCCA
 CAGTGACAAGTCCTACTGCTACTGCATCTGTCAATTATGTCTGAAACCCAACGAACAA
 GATCCAAAGAAGCAAAAGACCAAATAAAGGGCCTCGAAGAACAGAAACAAACAGC
 AAACACCACCCCCCAGGCAGGTTCTGGCTATACTGCATACTCAGCTCTAACAAACAGC
 TGATACCCCCCTGGCTTCAGTCATTCCCCACGACAAGATGATGGTGGAAATGTAAG
 TGCAGITGCTTATCACTCAACAACCTCTTCTGGCCATAACTGAACGTGTTGAGAA
 GTACACCCAGACTTTGGAAATAACACAGCTTGGAAACACAGTGTGAGCAAAT
 CACAGGAGAGTACCACTGAAAAGAGCCTCAGACACACCACCACTCCTCAGC
 AGTGGGGCGCCCCAGTGCCCACCTCCTCCCCACCTCCTTACTAAGGGTGTGGTT
 ACAGACAGCAAAGTCACATCAGCTTCCAGATGACGTCAAATAGAGTGGTCACCAT
 ATATGAATCTCAAGGCACAATACAGATCTGCAGCAACCCCTCAGCAGAGGGTAGCC
 CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTCCATCCAC
 TTCTGTGCCAGCACTAAGGTAGATAAACACAGAATTCTAAATGGAAGCCCTCTCC
 CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAACGGCA
 AAAGGCCAGCAGTAAGCATGCCCCCACCTCAGCCTCCAGAGGCCAGCACTCAT
 GCCTCACACTGGAATACACAGAACGATGCAGAAAGAGTGTGTTGATAAGAACCC
 TGGTCAAACCCAACTCCAAACATCTGCCCTACGTCTCTACCTAACAGACTCTATTG
 AAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTACAGTCCAGCTAAC
 AGACGTTTCTCCTGTGAGGCTGTGGAGACCCACTGCCCATCATCCACTGGACC
 AGAGTTTCACTCAGGANTGAAATATCCAAGGGACACAGAAAAGCCGTTCCACGT
 GCTTCCCAATGGCACCTGTCCATCCAGAGGGTCAGTATTCAAGGACCTGGAGAGTA
 CCTGTGCTCTGCATTAACTCCACTGGCGTAGACCATTTCATGTCCTTGTG
 GTTTTTACCCGGCAAGGATTGGAACAGACATGTCAAGGAGATCACAGITCACTTT
 GGAAGTACTGTGGAACTAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC
 CTGGATACTTGCAACCAAACGGTGGCTCAGAAACGGCAAGGAAAGCAGAAAGG
 TCTGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTATGATGTG
 GTTTTACAAGTGTGAGGCTGTGGCCAGCAACCCATCTGCCAGGATTCACTGTTG
 TACAAGTCATCACAGCTCCCCCTGTCAATTAGAGCAAAGAGGCAAGCCATGTTG
 GGGTTTAGGTGGAAGTTGAAACTGCCCTGCACTGCAAAGGAACCTCCAGCTA
 GTGTTCACTGGGTCTTATGATGGGACTGAACCTAAACATTGCAAGTGTACTCATT
 CCAGATTCTCTGTATCCAATGGAACCTCTGTATATAAGAACATCGCTCCCTCAGT
 GAGGGCACTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAACGGTAG
 TGATTCTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAAACTGCCCTCTCAG
 AAATGGACTGAGGTGAATTGGGTGAGAAATTACTACTGAACGTGCTCAGCTACTGG
 GGATCCAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCACTGACCAAGT
 GGCACAGAATGGCAGCCGAATCCACGTACCCAAATGGATCCTGGTGGGG
 TCAGTGACGGAAAAGACGCTGGTACTACTTATGTGTGGCAAGAACAAATGGG
 AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAATTGAACA
 GAAGCAGTATTAAAGAACAGTGCCTCATGGAAAGATTCCAAGTTGACTGCA

AGGCCTCTGGCTCCCTGTGCCTGAGGTATCCTGGAGTTGCCTGATGGGACAGTGC
 TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCCTT
 TTCCACAATGGAACCTTGTATTCAACAACGTTGGGATGGCAGAGGAAGGGAGATTAT
 ATCTGCTCTGCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT
 TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG
 CTGGAGAAACAGCTGCCTTGACTGCGAGGTCACTGGGAACCGAAGCCAAATGTA
 TTTGGTTGCTGCCTTCAACAATGTCATTCTCAATGACAGGTTCACATTTC
 ATGCCAATAGAACCTTGTCCATCCATAAAGTGAAACCACCTGACTCTGGGACTATG
 TGTGCGTAGCTCAGAACCTAGTGGGATGACACTAACAGACATACAAACTGGACATT
 GTCTCTAAACCTCCATTAAATCAATGGCCTGTATGCAAACAAAGACTGTTATTAAAGCC
 ACAGCCATTGGCACTCCAAAAAAACTTGTACTGCAAGCAGATGGGATCCCATCT
 TCCCAGGTACGTGGATTATGCCAGGCAATATTTCTCCAGCTCCATACTTGG
 AGCAGAGTCACGGTCCATCCAAATGGAACCTGGAGATGAGGAACATCCGGCTTC
 TGACTCTGCGGACTTCACCTGTGTGGTCCGGAGCGAGGGAGGAGAGAGTGTGTTGG
 AGTGCAGTTAGAAGTCTAGAAATGCTGAGAACACCAACATTCAAGAAACCCATTCA
 ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCGTAGCACTGAAC TGCTCTGGAT
 GGGAACCCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTGCTAAC
 AGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCATCCTTACAAA
 GCAACTCGAACAGTCAGGGAAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA
 CATCGAGAAACTCATCCTGTTAGAGATTGGCAGAACGCCAGTCATTGACATACGA
 ACCAGGGATGGTGAAGAGCGTCAGTGGGAACCGTTATCACTGCATTGTGTC
 ATGGGATCCCAAGCAAATGTCAGTGGACTACACCGGGTGGCCATGTAATCGAC
 AGGCCTCAAGTGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTC
 AGCAACAAACAGTCACGCCAACAGAAATTATCTGTAGGGCTAAACAGTGTG
 GCCAGGCAGTTATTAGCGTGTCACTGATGGTTGGCCTACCCCTCCGAATCATAA
 ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT
 GTGGCCTGGGAATCCCAAGCCAAAGTCACCTGGAGACGCCAACAGACTCC
 GCTCTAAAAGCAACAGCAAGAAAACCCATAGAAGTGGAGATGCTTCACCCACAAG
 GTACGCTGGTCATTCAAGAACCTCGGATTCCGGAGTCTATAAGTGCAGAG
 CTCAGAACCTACTTGGGACTGATTACGCAACAACTACATCCAGGTACTCTGACAGG
 AAGGGGGAGACTAAAATTCAACAGAACGCTCACATCCACAGGGTTATTTTGGAA
 GAAGTTAACAAAGGCAGCCATAGGCATGTAATGAGTCTGAATACATTACAGT
 ATTAAATTACAATGGACATGCGATGAGACTTGTAAATGAAAGCATTGTGA
 AACCGAGTCTCTGTGGATCTCAAAGCAAACCTTAACCTAAGGCACATTGATTG
 CAACAAATAATAACAAACATTAAGAGAAAAAAATGATCCACTACGAAATAACAAAC
 GGCTAATGCACCTGAATTCTCAGTAAAAGACCTTCTCGCTAACAGTTGCCAGC
 TGCCTCGTGTCTGTTCTACCAATGTACAAACATCGCACACAGGGTGAATGGAGT
 CAACGGGAAAGATTAAGTTGCAGGTCTGTGAAATCTCAATGTACAAATATTCTGTC
 NCTGGTTATAAACATTGATAAAACGAAAAA
 AAAAA (SEQ ID NO: 1)

Figure 2

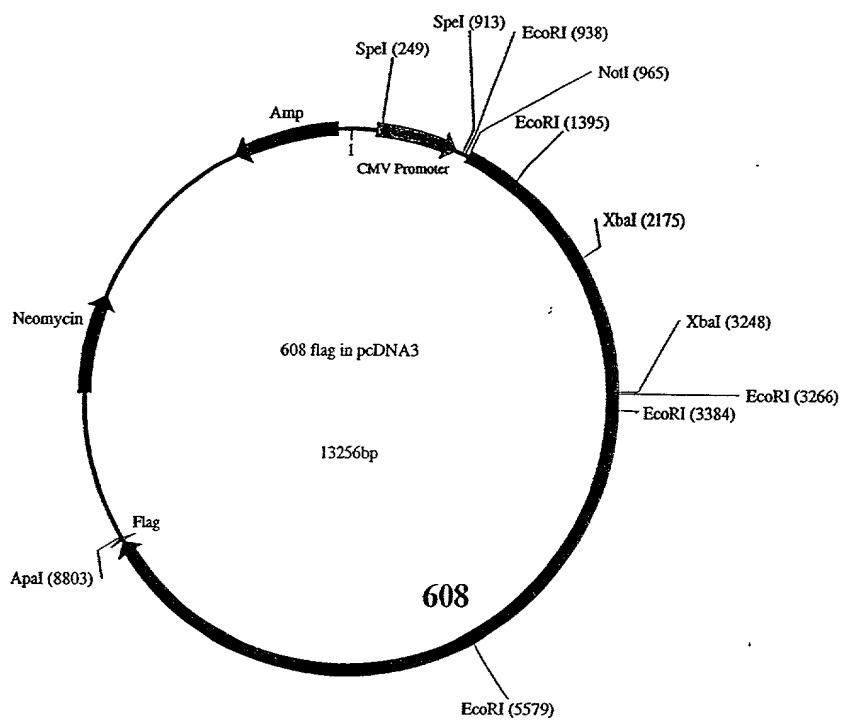


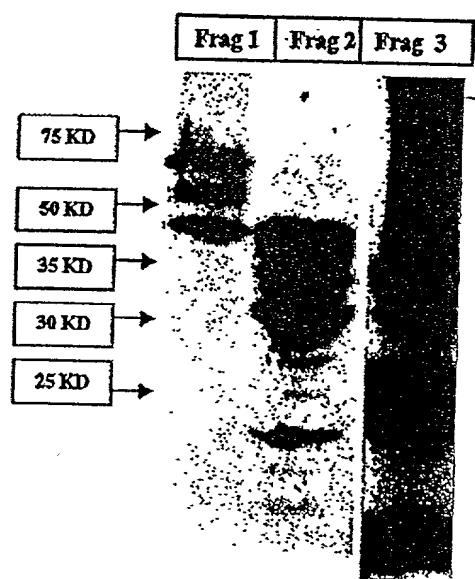
Figure 3

MQVRGREVSGLLISLTA VCLVVTPGSACPRRCACYVPTEVHCTFRYLT SIPDGIPANVE
 RINLGYNSLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFGLQLVKMSYNKVQII
 RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSRLVHLEGNRLLKHPDTFVSLSYLQI
 FKTSFIKYLFSDNFTSLPKEMVSYMPNLESLYLHGNPWTCDCHLKWLSSEWMQGNPDI
 IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSLSVLTQEDNGS
 ASTSPQDFIEPFGSLSLNMTXXSGNADMVCASIQQPSRTSPTAFTEENDYIMLNASFSTNL
 VCSVVDYNIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
 ADPFWFQQEKIVLQLNRTATLSTLQIQLFSTDQAQIALPRAEMRAERLKWTMILMMNNPK
 LERTVUVGGTIALSCPGKGDPSPHLEWLLADGSVKRAPHVSEDGRILIDKNGKLELQMA
 DSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
 PDASISWILPGNTVFSQPSRDRQLNNGLRQLQVTPKDQGHYQCVAAAPSGADFSSFKV
 SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPSLCLSALTGSEAGKQVSGVHRK
 NKHRDLIHRGGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
 KPVPLAVPLVELTDEEKDASGMIPPDEEFMVLKTKASGVPGRSPTADSGPVNHGFMTSI
 ASGTEVSTVNPQTLQSEHPDFKLFSTVNTGTA VTKSMNPSIASKIEDTTNQNPIIFPSVAE
 IRDSAQAGRASSQSAHPVTGGNMATYGHNTYSSFTSKASTVLPQINPTEYGPQIPITGV
 SRPSSSDISSHTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGV
 KNPHTPVLRRRHRHTVRPAIKGPANKNVSQVPALEYPGMCHTCPSAEGLTVATAALS
 VPSSSHSALPKTNNGVIAEESTTVVKPPLLFKDKQNVDIEIITTTKYSGGESNHVIPTE
 ASMTSAPTS VSLGKSPVDSNGHLSMPGTIQTGKDSVETTPLPSPLSTPSIPTSTKFSKRKTP
 LHQIFVNNQKKEGMLKNPYQFGLQKNPAAKLPKIAPLLPTGQSSPSDSTLLTSPPPALST
 TMAATQNKGTEVVSGARSLSAGKKQPFNTSSPVLPTISKRSNLTNFLSTETPTVTSPTAT
 ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGSYAISALTTADTPLAFSHSP
 RQDDGGNVSAVYHSTSLLAITEFEKYTQTLGNTTALETTLLSKSQESTTVKRASTDTP
 PPLSSGAPPVPTSPPPPFTKGVVTDSKVTSQAFQMTSNR VVTIYESSRHNTDLQQPSAEAS
 PNPEITGTTDPSNLFPSTSVPALRVDKPQNSWKPSPWPEHKYQLKSYSETIEKGKRP
 VSMSPHSLPEASTHASHWNTQKHAEKSVDKKPGQNPSTSCHKLPYVSLPKTLLKKPRIG
 GKAASFTVPANSVDVFLPCEAVGDPLIHWTRVSSGXEISQGTQKSRFHVLPGNTLSIQRV
 SIQDRGQYLCASFNPLGVDFHVSLSVVFYPARILDRHVKETVHFGSTVELKCRVEGMP
 RPTVSWILANQTVVSETAKGSRKVWVTPDGLIYNSLYDRGFYKCVASNPSGQDSLL
 VKIQVITAPPVIIEQKRQAIVGVLGGLKLPCTAKGTPQPSVHWVLYDGTELKPLQLTHS
 RFFLYPNGLYIRSIAPSVRGTYECIATSSSGSERRVILTVEEGETIPRIETASQKWTENV
 LGEKLLLNCATGDPKPRIWRLPSKAVIDQWHRMGSRIHVYPNGSLVVGSTEKDAGD
 YLCVARNKMGDDLVLMHVRLRLTPAKIEQKQYFKKQVLHGKDFQVDCASGSPVPEV
 SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFNNVGMMAEGDYICSAQNTLGK
 DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS
 NDRFTFHANRTLSIHKVKPLDSDGYVCVAQNPSSGDDTKTYKLDIVSKPPLINGLYANKT
 VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMNRNIRLS

DSADFTCVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVGDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLLYKATRNKSGKYRCAARNKVGYIEKLIL
LEIGQKPVILTYEPGMVKVSGEPLSLHCVSDGIPKPNVKWTPGGHVIDRPQVDGKYIL
HENGTIVIKATTAHDQGNYICRAQNSVGQAVISVSVMVA YPPRIINYLPRNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLGTDYATTYIQVL (SEQ ID NO: 2)

8/90

Figure 4



9/90
Figure 5

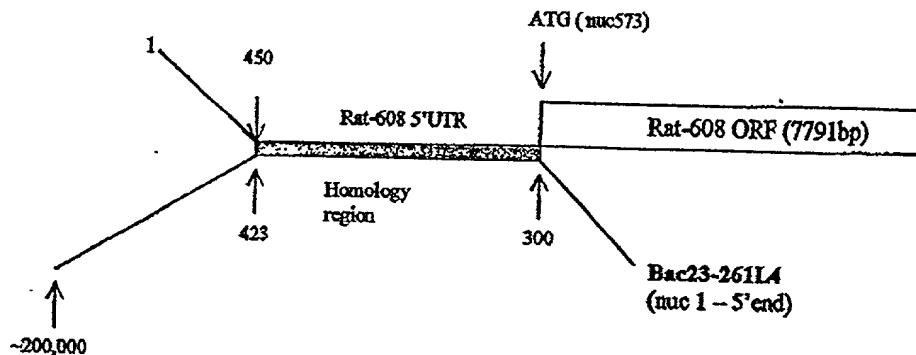


Figure 6

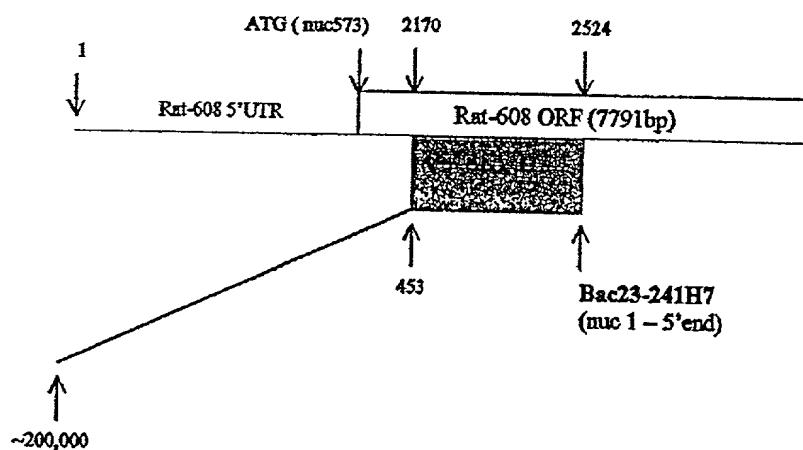


Figure 7

Nuc 1
 ↓
 TTTGGAACCAACCCAGATGCCCTAACAGAGAAATGGCCAGAAAATGGTCCATTAA
 TCCAATGGAATACTACTCAACTTATTAAAACAACGACTTCAAAATTTAGGCAAAT
 GNAATGGTCTGNAGGATCTTGAATGGAGGAAACCCAACTCAAAAAGAACACTCATGGTATGC
 ACTCACTGATAAGTGGCTATTGGTCTATGGTGTAAAAGGGAAAGAACACATAGC
 TTTTGTGTATAATATTAAGATGGAAATTGCCAGTGTGTGTTGGCTATGAGTGAATCT
 TGTTTCACTGGATTACCGGAAGAAAATAATAAGTGAACGTAGGAAGAAGTAGTAAATCA
 ↓ Nuc 390
 AGGTGACAAAGTATCCTGACACATTGGGAAAGACCCACAGTCCAGGAAACTGAGTCTAA
 GGATTCTATTAACCTCAGTCCCTACAGCAGTCTGAGACTTTGGCAGATCAGACACTT
 AACCTCACCCTCTACACAGCAGTCTACTTCTCTGCACTCACACATGGAGTGTGA
 CCATTAAAGCTGACGAAACATGAGTCTGACTTGTATAATCTTAAATACAAATGTTG
 TGTAAGTATGTAACCAAGACATGGTCACTGCTATTAAACCTTGTGATGTTCTATGGACCT
 TAATTTATGGTAGAAATGTCACAAGCTTGTGGAGGCTGGAGATAACAAGGCTAAAGAG
 GATGGCCTTCTGAGTTTGAAGTAACTCAGTATGTTCTGGCATOCCCTTCTAAAGCA
 ATTTAACCCCCAAGTGGCATAATTAAATGCTTACTCTCAGAATATGCTAAATGAC
 TCTCTAAACGACTTGGTATGGTCACTGTTAAATGTAATATGTAATTAATGACATAAT
 AAATAGGAGAAACGACTGAGCTAGAATAGGTTAAATATGTCGCTGGCTTCTAATAGGCTT
 ATAGGTTATAAAAGGGGGAAAGGAATATTGAAACATCTAGAAGTAAATGATCCTG
 AGTAGCGATCTGGGAAATACTGACTCTAACACACTGCAATCATCTCTGTTGCT
 GGAGCTGAGGTCTGGAAAGGCTGACCTTGGTAAAGGAAATAACCTACCGAATACAGAGCTAT
 GAOGTAGTCTGGAAAGGAGCTTGGAAAGAATGACAAGCTGAGCTGCOCCAGAACATACTA
 GATGCCATATTCTCAAGGCAAGTGTGCACTGGGAGCATCTTAAAGATAATGTCCTG
 CAGTGTAAAGGACCTGTCGACACAGGCTCTCCAGGGTGTGCTAACCTGACTG
 TTGACTCTTAACTTACCTGATCATTAAATGACTAGAAATCACTTGTGATTAGCAACTG
 GATATGGAATATTACTAATTGTAACCCAAAGCAGGGCCTACTCAGCTTGGCAGCTCATTG
 ATTCTGTTGGAGGCCAGTCACGGGGTTGAACTCAACTGACTCTTCTACTTACAAGACGC
 ATTACCTGAGATGAGTCACTTCTCACAAGCTTCTTAAAGACTGAAATGACATATT
 CTGATGAGTAAAGCATATAAAAGTGGAGAGGAGCATGAAATGCTTCCATGATGCTCATGGA
 TGCTATTATAATGTTGAAATAACTGACTTTAAACGCTTATGATACTTGTCA
 GAGTAAATCTCCATAAAATCATCTGCACTTATAAAATTATTTCTATAATCCATCAATTAA
 AAACCTTGTGAAATTGTTAACACAAAGACCTAGGCCCCCTGGCCCTAGGATGGTCTGTA
 TGGTGGCTGGAGAGGAGCTTAAAGACTTACTGCTCCAGGGACACATCTTCAGAAC
 ATCTGCTCAAAACATTATCCTAAATGCTCATCAAAGGCTCACTCACATGTCCTCAACC
 ACAGGGATTAAACAGTCATTGTCACATTCTCAACCGGCTGAGGCTGCTAGAGGAA
 CAGGATGTTACGGATAACATCCAACCTTACAAGGGATGTCATAACCTCACCACAA
 AACAAACACGACAACAAACCCATAAAATTATCACGGCAATGAACTAAGGCAATGCA
 GAAAAGTATTATGTCATGTTGGGTTCTTCTTAAATGTCAAATATGCAAGAAT
 ATAGACAAGGATGTTTATGCAAGTGGGATGGCAAGGAGATACTTGTGATTAGAGGACA
 CAAGGCAACAAACTACAGAGTGAAGTAATCCAGAGACTTAAATGTAATATGAGGACCTA
 TTTAATAATTCTATTAAAGATAACAGCAAACGAGTGTATCTTACTAACACACACCTAC
 ATAGAGAGATAAAAGTGTAGATACTGTTGTTCTCTCATGTCAGCTGATAATTCTATAT
 TGACACCTCAACATGATAACCAACAAAGAGGGAAAGGAGTATGGTGCCTCTCCAGGGC
 GGAAGAGTACATTGAAAGTCAGACACCCATTGTTAGATGTCACCCATGGAGGGAGCTAGA
 GAAAGTACGCAAGGAGCTAAAGGGATCTGCAACCCATAAGGTGGAAACAACATTATGAGCT
 AACCACTACCCCGGGAGCTCTGACTCTAGTCAGCTATATATACAAAAGATGGCTAAATGG
 CCATCACTGGAAAGAGGGCCATTGGACTTGCACAAACTTATATGCGGCTACAGGGGA
 ATACCCAGGGCCAAAAGGGGGAGTGGGTGGGCAAGGGGAGTGGGGTGGGATATGGG
 GGACTTGTGATAGCATGGAAATGTAATGAGTAAATACCTAATAAAATGGAAAA
 ↑ Nuc 2817
 AAAAAAAAAAAAAAAAAAGGAAGGTCAAGACACCTCACTTCAGTCTATCTCAACTT
 GCAAAACAGAAGGGGAGTCACAAACCCAGGACAAACACACAGTGTGAAAGCGTCTTGA
 TGTTATTGCTGTTGTTGTTACCCATCATTAGCATATATTCTTGTGAAACCTACGGGGT
 CTATGACATGTTTCTTATCAAGTATATCACATGTCAGCATACTGGCACCACCTACCA

Nuc 3114

▼ Exon 1 (Nuc 3661)

GCTGGGGTCAGGGCAGACAGAGGGTGGGATGGCAAGGGTCCAGGAAAAGGAAAGT
ACTGGGGGGAGTTGGACAAAAGCAGOGACCAAGGGACATCGCTTCAGTGAAGCCA
GGCAAAAGGAGGGGAAGGATTATAATGAGCCTGGGAOGCTTCTATAAACACTGTGAGCTGT
TTGTCAGAACAGCAATTGGAGGAAAAGCCTGGGAGCTGGAAAGGAGGAGTGTGCAATT
AGTACTTGTAAAGTTAGGTGAGTTGAGAACAATCAACTAACCTATACATTAGGGAGAAGGAA
AGCATTCCAGCAGCAGCAGCAGCAATCAGATAAGGAAACCTTGGTAAAGTGTGGAA
ATGTATGATAACATTAAATAACAGAAGOGCTCAGITCTCTGAAAGACTCAGTCCCCAGCTA
GTGAAGACTAAGCCTAACAGCTTGTCCCGTTGGAAAGCAAGAAAGCCTTCAATCAGG
TGAAGGCTCTCTGAGAAGATTCTGTCTGCTATGTTACAAGAGGATTCAAAAGCAAGAC
AGAAGCTCAGTATGCCAATCTT

4 Nec 4203

TGTAAATACAGTTGGCTTAAGTGTACGGGAACCTCATGGTATTCACTTACGGCTCT
CTTCTCTTATAACTAACTCTTAAGGTGCATATAGCTCTCTCTGTTCCAGCTACCTTGAC
CACTTTGTTATAATAGCAAGCTCATGCTTTAATCARCAAGCAGAGAGTATT
AAAAATATTCAGTGTAAACAGTGACAGTGAGGCTAGAAGTAATCTTAAATC
TTAATTGGGTTAAACTCATCTATAACGCTCAGGTTGGGAGGATCAGTGAGCTTGC
CAGCTGGGGTTAAAGATTTCTAACAAAGAGAAGCAGAATTCTCTTGCGATGCTC
CCATCACTGTGTCAGTAAGCAGAGGGGTGTTCCAAGCAGAGAAGAGCAGACAGTGTTA
TGCTGCAAAAGTCAGAGACTCAGCCTCCAGCTGGTCAGTTACTGTCTCCGGTCATT
AGTGGCTGCAAGGAAAGGCCATGTCCTTATGGCAAGGACTTGCAAGACATGCTAGAAA
GAAAATTGGACCTTTTCTAGTGGGTTTACAGCTGTTAAAGTTTGGAGGGTTAAAG
CCCCAATAAATAAAACACATATTAAATATACAAATGTTACAAAATTGATCATATAAGAA
GTACATTCAAAATGCAATGTTAAAGATATATATAATTATATCTTATGTCAGGAAAG
TTTCTAAATGGCACATGTAATTTTATTTATTTAAATTTAAATTTAAATGTTATATAAA
AGTGTAAAGGCTCTGGCTAACAAAGAGAGTTAAATCAAACTTTAATGTCGGGAT
AACATTATATGCTTAATGACAGGGATCTGCTTCTGTTGGAATGGAGAAGCTATG
AGAGATTTACATAATAAAGGCCATTAGTGTAAAGTCCAATGGGAAGCTGACAC
ACIGGTTATAAAGAGAACAGTTCTGAGTCATGCAAGTTACACTCTAGGGAAATAAG
AGTTCCTCTTCTCAGATTCACTAGCAATTGTTGTCATCATTTACTCTCTGATGAG
CATTATAAGTGGAAATAAGATAGGATCTAAAGGAATGTCATTGGATGCCCTGAAACAAT
CTTCTAGGGCTCTCTCAGTCTACTAGCTTATGTTGATAATTGGGGATGGGTT
TAATTCTTCTCAGTTCTATTGAAATTCCAAAAAAACAAAAACAAAAACAAAAA
ACCTCTGAAACTAGAACCTAACCTACATTACTGTTGTTGAAACAGGAAATCTGCA
GAATTATGCTACATTGTTCAATTTCAGACAGGCCAGAAATGTTGAAACCAACTTACGIA
GCCGTCAAAATATGAAGGATAAAAGAAATGTTGAAATGTTGACACAGAGTOCCATG
GCCATAAAGAGTGAATCATGACATATGCAAGGAAATGGATGCAACTGGAAATCAATTG
GGCTAATCAAAACAGACAGACTCAAAGGAAACCCGGTGTAGCTCTCTGACAAACA
GAAGCTGAGTAACTCTGTCAGTGCGCATGTTGTTGAAATTATTTAGTAACTACT
ATTCTAATCTGAGTGTGATAAAAGGCATGTCATGTAAGCAAAAACAACTGAGCTGGGG
TOGGTAGGAGAGAAAGCAATGAGAGGAGTTAAAGAAGCAAGCATAGTAACATAGGIG
CCAGGATGAAATGCAATTAACTTGTATGCTAACTAAACCACAGACAGGAGGACACGTTCA
AAACAGGGTGAATACTCCAGCAGACAGAAGGGGAAGTAGACACAAAGTTGCGGCACTAAC
CAAGAAGGCTATTGCTAGGTTGCTGCCGCTGGGGAGGGGTTCTCAGTTCTCTGAC
ACTGTGTTATAACAAACCAAGTTGACAATACAAAGTGTGTCATGTTGATGTTGTTGTTG
TTCTATTCTTCTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
ATTGTTCTTCTTCTGAGAGAGAGAAGGAACATGAAATTGGGTGGTAGGAAGCTGGAAACG
ATCTGGAAGAAGGTTGGGAAAGAGAAAATGTTGTTGAGGCTATTAAACAAACAA
ACAAACAAAAGGTTCTATTGCTGCCAAACAAAGGTGTTGAAATTAAACCACTGAGCT
CTAAAGAAAATATCCAAATTATGCCAGAGTGTGTCATGTTGCTGTCCTAGGGCTTGTG

▼ Exam 2 (Nuc 6559)

CTTGAACCTGGCCCTATAACCTCTGGTGTTGCTTTCAAGGATGCAAGAGAGGGCAGGGAA
GTCACTGCTGTTGCTGATCTCCCTCACTGCGCATCTGCTGGCTGGTCACACCCCTGGAGGCAAGGGTC
TGTCTCTCGCGGATGTGCTGCTGCTATGGTGCCTGGCACTGTCATATTGCGGACCTGACCC
TCCATCCACGAGGGCATCCCGACCAATGTGGAAAGCTGCAATTAGGGTGTGT

Nuc 6755

GGACCTTGCTGATCTCCCTCTCAGAGAGGGACCACTGATTTCTGTTACTCTGGCCCC
AAACACCTGTGATTACTTTAATAGTTCTCTAAATGGGTICATACAAACCTTATATG
TGGAGACAATGAACATTATGCCATAGTCTTTACTAGAACTTGAGGCCCTCTTAGIT
GTTGGGAGGCCCTCATAAATTATGGGCAGCTTATCTGATGAATTTAAATGAAAAAGAT
ACAGTTCTGTTAACATCATTATGATACCAAGGAAGGAAATGTCATTGATATTITAA
AAAAGCATTCTTGTCAATTATATAACCCATTACAAAATGGCTTACTTAAATATCTG
CTTACTAAATCTGACAAATTATGGTGTATTTGAGGTTATGAAAATTGTTATGTTG
ATAAAATGCAAGAAAATGGGATATGCCATCACCTATGTGCCATTAGTGAGCATGTACAGT
ATGCCAACACTATTGTCACGGTGGAGGAAGTAATGGGGAGCAACAAGGGT
TATAACCGTATACCCAGTGCCTGGAAAGCGATTGCAAAACAGTAAAGACTGACATTGTT
CTCCCTATGAGGGAGGGGCTTGGCTGAGCACTTGTCAATGAGCATTGCTCATTTGCT
GGCAGGTTTATGATAACCTGACCCAAAGCTAGAGTCATGGAGAGGAAGAACCTCAACT
GAGAACATGCCATGAGAACAGATCAGATTATAGGCAGGCCGTGGGGCATTTCTTAATTAG
TGATTCTGGGGCAGGGCCAGTCCATTGTCGTGTTACCATTTCTCAGGCATTATAAAA
AAAAAAAAACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGGCCCTAATGA
TCTCTGTCATCAGCTCTGCCCTCAGGTTCTACCCCTATTGAGTTCTGCTCTAGCTCCCTA
CACTGATGAACAAATGATGTCAGGAAAGTATAAGCCAAATAACCTTCTGCCAACITGCTG
TTGGTCTAGATGTTCTACAGTGTGATAATAGTCCTCATGAAGATGCTGGTTTATAAC
CTCTTGGACTAAATCTGTTATCTATAGTCAGGAAAATGGAGCATAGAAATGCTCCAGA
CTACACCAAGAGTGTAAATCTGGGCCCTGAGCTAGAATCACCCACAGTGCACCTCAGTGC
GGGGCTCTTAAACCGGAACACAGTGTAAAAGGGAATTCTGTTTCTCATTGACA

Exo 3 (Nuc 8089) ↓

TGIGGACTTAATTGACGATTCTGAAAGCTGAAAAATGATTTTTTCCAGGTATAACAGC
CCTACATGAGTGACAGAAAATGACTTCTGGCTCTGACGAGACTGGAGTTACTCTGCTGCA
GAAATGGCATTCAAGACTGACAAAGACCTCTGGCTCTGAGCTCTGAGCAGTCAG

Nuc 8218 4

GTGAGATAGGTAGAGGGTGTGGAGGCTGAQAAAGAGAGGTCAACTGTGGGTATACCC
AAAAGCTCTGATTCCCGTGGGAGACATTCTATAAGCATCTATAAAACTAGAGGCAGATA
TCAAGGAAGGATTCAATTGTAATGCAATTCTATAGAGAAAATTGAGATAATTAGAAAATG
CTGGGAAATTAGCTTACACAATGGCAGGACCTAATTAGGATCTCCTAACATGGCACAATAA
AAAGCACACCATGGGGCAGACACCTGCAATTCTGTGGCTGAAGGCACCTGTCAGAAT
CCAGAGACTCATGGGCAAACACTCTATTCAATCAATGAAGTOCATATTCACTGTGACAAA
ACTTGACTCAGAAACTAATGTGGAAGGCATCAGGAAGACAGGCAACACTCTGGTCTCTACT
CATGCATGAATAAGGGATCCAGAGAGAAGGGAGAAAAGGAAGGAGGAAGGAGGAAGG
AAGGAAGGAAGGAGGAAGGAGAAGGAAGGAGAAGGAGAGGGAGGAGGAAGGGAGGG
AAGGAAGGAAAGGGAAAGGAAAGGAGATGGGGAGGGAGGGAGGAAGGAGGGAGGG
GGAGAAAGAAGAGAAGAAGGAAAATAATAAATTTCTAGGGATTATTACACCTTTAAA
TTTATCCATAAAAGGTCATTCCACCTGTTGTCAGGAAGTAGAGTGGGATCCCTTATAT
AAGGGCAGCTTCAACATAGTAGGACTTTATAAACCATTACAAATTGAGGTTCTCTCTAC
TTTATCTCTACCATCTCAAACTGAAACTACAAATTCTCCACAAATGAGAAGAAAATGC
TGTAGAGTTTACACACCGAAAGTGGGAAACTTAAGGATTAGACAAGTCTAACATGAG
AAAGGGGAGAACAAAAGAGACTGCACAGGGAGGCCTTCTGCTTAAATCTGACAC
TGTAGAAGCTAATTGACGCTGCGATGACTACTCACTCTTAAAGCAAAACATTCTGTTGTT
ATGAAAAGCACAATAAAAGTACATATGCTCCATAATATTCATCCTAAATTGCGATCAGCAG
ATAATAGCAATCAAGCAATAAACACCOACTGTTCACAGAGACTTTAAACATGAAACTGGA
ACTATGCTAGTGTGTTGACTTAGGGTACATAGTAGTGTGCTGCTGTGATGACCAATGTTG
ATTAGGTCTACAGACAGCATTGGAAACATGTATCTCAGGGAGGAATCATCTGATCT
GCTGATAAACTCTCCACCTATGTTTATCTCTTAAAGCCAGGTTTCTCTGATGAGGAAACATT
GGGTGAGGTTTACTCCAGGTAACATTAGGGAAAAGCTGTCTATGTTCTCAGTTGG
CTTITATITATGAGGGATTTGGTTATCTCAGAAAATTCTCTTGTGAGGATTACAAATTAA
GGTCAAAACAGAAAATATGAAAAGTTATTGTTTATTAGTATTCATGTTCTTTCT
TTTAAATGGTATGCTAGAACTAATTAAAGATTAGATTAGATTAGAAAATAATC
AGAGAGGGATTGATGAAATGCTAAAGCATCATGAAAATTCTCCTAAATTGCTCTAAAT
CAGAATCAATTAAATTCTATTAACATGAGGAGCAGGAGTGTGCGCAGCTGAG
GAGTGTGAAACCTGTAGCAGTGTGAGTGTCTATGAGAAAAGGAGTGAAGGGTGA

TGTGCGCTGCAACATCTGAAAACATTGGCTACATGATGGAAGGCCAGGCACAAAAAGCC
 ACATATIGCATGGGTTATGTTATATGAAATGTTAAAATACATGGATCTTAGCAGAACAGA
 GTAAAGATGTTACTTAGGGTCAGGAAAAGATTAAAAAAACTATTGATGTTGAAATG
 ATCTTAATTGGGAAAAGACAATTCTAAGACGAAATGTTGAGGTAGATAGTTAT
 ATCCCTGTTGATATTGTAATAAACCAAGCATGCTGCTGCTGAGAAGGGCTAATGAGGG
 GCAGGAGGAAGTGAATGAGATGGTAGAAAGGAAAGTCATATAACATGGCTCTGCTG
 GGTGGAACTAGATATGTTATATGACATAAAGGAGGAATGTTAGGGAGGATC
 AAAACCAACAGGAGTGAAGGGAGACATAGGAACCAATGAGAGGGCAAAGTICATGGTCAA
 TGTGTGAGACACATAATAAAACCTTTGCTAAGAATTTGCTAATCTAAACCACTAAACAT
 TAAAACAAAACATTTCACAAGAATATTATATTCAATAAAAGATGTTAAATGGGG
 GAAGTGAAGTICATTGATAGTCTCATATAATCTAAATGTTAAACCTGCTTACGTT
 TTATTATTAATTACTCTGCTGTCATTATTATCATCATTATGTCATCATCATCACTA
 ATGCTTTTACCCATACACAAATGTAAGGAGAAGTCAATTCACCTAGTGAGGCAATCTT
 GGAGAGGGAAAGGAAGCGGATCGGGGAGAGGCCACACAGGAGGACAGTGAGAGGGAA
 ATGAAACAAGAAAAAAATGTTGACACATGCACTAAACCACTAGTCCACTACATTACTTG
 TATTCTAATATAAGAAAATAATAAAACCCATTCGTGCACTTATCACCAGGCTAACAG
 TTATCTGCCACAGATCCTGTCATGCACTGTCACCTGAGTCCACTTACGCTCTG
 AATCCAATCCAGGGCATGCTTACTCTACACAGAACTAAAGATAAAGAGAGTTAA
 AAGTAACCATGACATCTCTGTTCTTACGCGATAAGTCTTAATTTATGGCTGCTTGT
 GTATGTTCTAATTCTCTAATATTGTCACATTTAGTGGCAACTACTTGTGTTGAATTGAGT
 TGGAGTTAAGGTCCTAGGATAATCTCAACATATTCTATTTATAAACTTTCTCTCT

Exon 4 ? (Nuc 11289) ↓

TTGTGAAAGTTCTTGTGAGAAAACAAATATGCCATATCTTCTTACAGGTCTAAAAATG
 AGCTATAACAAAGTCCAATAATTGAGAAGGATACTTGTATGGACTCAGGAGCTTGACCCGG
 TTGCACTTGGATCACAAACATTGAGTTTATCAACCCGGAGGGCTTACGGACTCACCTTGC
 TCCGCTTGGTACATCTAGAAGGAAACCCGGTCAACAGCTCCATCCAGACACATTGTCCTT
 GAGCTATCTCCAGATATTAAAAACCTCTTCTTCAATTAGNACCTGTACTTGTATGATAACTCATG
 ACCTCCCCCTCCAAAAGAAATGGCTCTCTATGCCAACCTAGAAAGCTTACTTGCGATGGAA
 ACCCATGGACCTGTGACTGOCATTAAAGTGGTTGTCOGAGTGGATGCAAGGGAAACCCAGGTA

Nuc 11680 ↑

ACTATCTTGTGTTGTTCTTCTTCTTATARKACGTATTCTCTCAATTCTTACATTAGAATGA
 TATCCAAAAGTCCCCATAACCTCCCCCACTTACCTACCTACCCATTCCATTCTTCTTGG
 CCTGGCATCCCCCTGACCTGGGCACTATAAAGTGTGCGTGTCCAAATGGACTCTCTCTTCC
 AGTGATGGCCAACTAGGOCATCTTGTATACATATGCACTAGAGTCAAGAATCTGGGG
 TACTGGTTAGTTCTATAATGTTGTTGCACTACAGGGTTGAA

Nuc 11967 ↑

(SEQ ID NO:3)

14/90

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

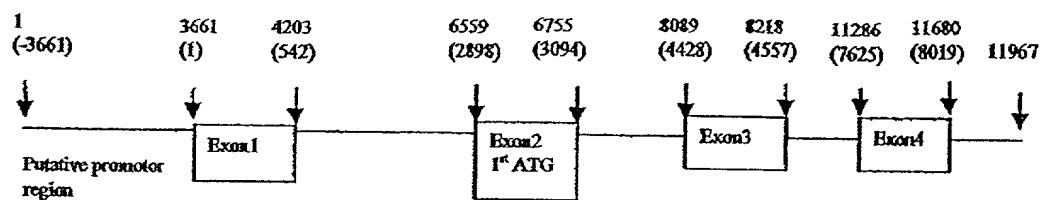


Figure 10

cDNA_rat	951	acggactcggagcttggtecggttgcacctggatcacaacaacattgaa
genomic_hu	1	-----
cDNA_rat	1001	ttcatcaaccctgaggcctttatggacttacctatccgttggtaca
genomic_hu	1	-----
cDNA_rat	1051	ttagaaggaaaccggctcacaaagctccatccagacacatttgtctat
genomic_hu	1	-----
cDNA_rat	1101	taagctatctccagatatttaaaacctcttcattaagtacatgttctt
genomic_hu	1	-----
cDNA_rat	1151	tctgataacttcctgacctccctccaaaagaatggtctctatgcc
genomic_hu	1	-----
cDNA_rat	1201	aaacotagaaagectgtatggatcatggaaacccatggacactgtgactgcc
genomic_hu	1	-----
cDNA_rat	1251	attnaaagtggttgtatggatggatgoagggaaaccoagatataaaaa
genomic_hu	1	-----
cDNA_rat	1301	tgcaagaaagacagaagatcttccatgtccatgtcccttgcatt
genomic_hu	1	-----
cDNA_rat	1351	gaaccccaggatctctaaaggcagaccctttgtatggtaaccatctggag
genomic_hu	1	-----
cDNA_rat	1401	ctttccatgtacaaagccaaaccattgtatccatcactgaagtgaaagago
genomic_hu	1	-----
cDNA_rat	1451	ctggttactcaggaggacaatggatctgcctccacctcacatcaagattt
genomic_hu	1	-----
cDNA_rat	1501	catagaaccccttggatcttgcattttgaacatgacananntctggaa
genomic_hu	1	-----
cDNA_rat	1551	ataaggccgacatggctgttagtacatccatcaaggacatccatca
genomic_hu	1	-----
cDNA_rat	1601	actgcattcactgaagaaaatgactacatcatgotaaatgcgttcatttc
genomic_hu	1	-----
cDNA_rat	1651	cacaatcttgcgtcagttgttagattataatcacatccagccagtgtggc
genomic_hu	1	-----
cDNA_rat	1701	aacttctggatatacagtgtactctcttotgtatactagaagaaagccc
genomic_hu	1	-----
cDNA_rat	1751	cagcttaccgagactcttcaactgtctttagatataaaaggtggatot
genomic_hu	1	-----
cDNA_rat	1801	taggcctgaagacattttaccagcatagaggctgtcagagccagacc
genomic_hu	1	-----
cDNA_rat	1851	ctttttggttccaacaagaaaaattgtcttgcagctgaacagaactgcc
genomic_hu	1	-----
cDNA_rat	1901	accacacttagoacattacagatccatgtttccactgtatgtcaatcg
genomic_hu	1	-----
cDNA_rat	1951	tttaccaaggcgagatgagagcggagagactcaaatggacoatgtcc
genomic_hu	1	-----

<i>cDNA_rat</i>	2001	tgtatgtacaatccaaaactggAACGCACTGACTGGTTGGCGGACT
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2051	attGCCOTGAGCTGTCAGGCAAAAGGOGACCCCTCACCTCAGTTGGAATG
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2101	gtttotAGTGTGGAGTAAAGTGGAGAGCCCTTAAGTACGGAGGATG
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2151	ggcGAATCCTAATAGACAAAAATGGGAGTTGAACTGCAGATGGCTGAC
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2201	agctttGATGCAAGGTOTTTACACTGCATAAGCACCAATGATGCAAGATGC
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2251	ggatGTTOTCACATACAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCA
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2301	cacatGACAGTGGAGTCCAGCACACAGTGGTTACGGGTGAGACGCTGAC
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2351	cttccatGCTTTCCACGGGTGTTCCAGATGTTTATTAGTGGATTCT
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2401	tccAGGGAACACTGTGTTCTCTAGCCATCAAGAGACAGGCAAATTCTTA
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2451	acaatGGGACCTTAAAGAATATTACAGGTTACGCCAAAAGATOAAGGTCT
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2501	taccatGTTGGCTGCCACCCATCAGGGCCGACTTTCCAGTTAA
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2551	agtttcAGTTCAAAAGAAAGGCCAAAGGATGGITGAGCATGACAGGGAGG
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2601	caggTGGATCTGGACTTGGAGAACCCAACTCCAGTGTTCCTTAAGCAG
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2651	coAGCATCTTGAAACTCTCTGACATCAGCTTGCAGGGTCAGGGCTGG
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2701	aaaacaAGTCTCGGTGTACATAGGAAGAACACATAGAGACTTAAC
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2751	atCGGCGGCGTGGGATTCCACGCTCGGCGATTCAAGGGAGCATAGGAG
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2801	cagTCCTCOTCTCTGCTGGAGATTGACCGCGAACGCTGGCAGCACT
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2851	tctAGAAAAGCCAAAAGAATTCTGTGCAAAAAAGCAAGAAAATACCA
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2901	cagtaaAGCCAGTGCCACTGGCTGTTCCCTCGTGGAACTCACTGACGAG
<i>genomic_hu</i>	1	-----
<i>cDNA_rat</i>	2951	gaaaAGGATGCCCTGGCATGATTCCCTCCAGATGAAGAAITCATGGTTCT
<i>genomic_hu</i>	1	-----

cDNA_rat	3001	gaaaactaaggctctgggtcaggaaaggtcaccaactgtctgactotg
genomic_hu	1	-----
cDNA_rat	3051	gaccaggtaaatcatggtttatgacogagtatagcttctggcacagaagtc
genomic_hu	1	-----
cDNA_rat	3101	tceactgtgaatccaaacactacaatctgagcaccttcttgcattcaa
genomic_hu	1	-----
cDNA_rat	3151	attatttttagtgaacaaacggtaacagctgtgacaaadagtatgaacccat
genomic_hu	1	-----
cDNA_rat	3201	coatagcaagcaaaatagaagataacaacaaacaaaacccaaatcattatc
genomic_hu	1	-----
cDNA_rat	3251	tttccatcgttagtgcatttgcatttgcaggcaggaaagcato
genomic_hu	1	-----
cDNA_rat	3301	ttcccaaagtgcacacocgttaacaggggaaacatggctacctatggcc
genomic_hu	1	-----
cDNA_rat	3351	ataccaaacacatatagtgcatttaccagcaagccatgtctgcag
genomic_hu	1	-----
cDNA_rat	3401	ccaataaatccaaacagaatgttatggacactcgatcttacaggat
genomic_hu	1	-----
cDNA_rat	3451	cagcagacacatgcgtgtgcacatcttctcacactactgcagacccta
genomic_hu	1	-----
cDNA_rat	3501	gtttctcoagtcacccttcagggtcacacaccactgcotgttttattt
genomic_hu	1	-----
cDNA_rat	3551	cacatttctagaacaaatacaggtaacttccottgtccaggactt
genomic_hu	1	-----
cDNA_rat	3601	gggaagagagaggacaatgtggagcagagggagaggtaaaaacccacata
genomic_hu	1	-----
cDNA_rat	3651	gaaccccaagtctccgacggcatagacacaggactgtgaggccagaatc
genomic_hu	1	-----
cDNA_rat	3701	aagggacotgctaacaaaatgtgagccaaatccacagacttacccacat
genomic_hu	1	-----
cDNA_rat	3751	tgggatgtgccacacatgtccttccgtcagaggggtcacoagtggotactg
genomic_hu	1	-----
cDNA_rat	3801	cagcaactgtcaggtaacatgttccatcccaatgtgcacacttcccaaaactaat
genomic_hu	1	-----
cDNA_rat	3851	aatgttgggtcatacgacaagagacttaccactgtggtaagaaacact
genomic_hu	1	-----
cDNA_rat	3901	gttactatataaggacaaacaaaatgttagatattgagataataacaacca
genomic_hu	1	-----
cDNA_rat	3951	ctacaasatattccggggggaaatgttacatgttccctggaaatctcctgtags
genomic_hu	1	-----
cDNA_rat	4001	agcatgacttctgtccaaatctgttatccctggaaatctcctgtags
genomic_hu	1	-----
cDNA_rat	4051	caatagtggcacctgagcatgcctgggaccatccaaactggaaagatt
genomic_hu	1	-----
cDNA_rat	4101	cagtgaaacaaacaccacttcccgccccctcagcacaccctcaatacc
genomic_hu	1	-----

cDNA_rat	4151	acaaggcacaaaattctcaagaggaaaactccottgcaccagatctttgt
genomic_hu	1	-----
cDNA_rat	4201	aaataaccagaagaaggaggggatgttaagaatccatataatcggtt
genomic_hu	1	-----
cDNA_rat	4251	tacaaaagaacccagccgcaaaagttccaaaatagctcccttttaccc
genomic_hu	1	-----
cDNA_rat	4301	acaggtcagagttccccctcagatttacaactctttgacaaggccg
genomic_hu	1	-----
cDNA_rat	4351	accagctctgtctacaacaaatggotgcactcagaacaaggcactgaag
genomic_hu	1	-----
cDNA_rat	4401	tagtatcaggtgccagaagtctctcagcagggaaagaagcagcccttcacc
genomic_hu	1	-----
cDNA_rat	4451	aacttccttcactgtgtttccatgcaccataagoaagagatctaatacatt
genomic_hu	1	-----
cDNA_rat	4501	aaacttcttgtaacggaaaccccacatgtacaactctactgtactg
genomic_hu	1	-----
cDNA_rat	4551	catctgtcattatgtctgaaacccaaacgaacaagatccaaagaagoaaaa
genomic_hu	1	-----
cDNA_rat	4601	gaccaaataaaggggctcggaaagaacagaaacaacgcacacaccacccc
genomic_hu	1	-----
cDNA_rat	4651	caggoaggtttctggctatagtgcataactcagtcataacaacagotgata
genomic_hu	1	-----
cDNA_rat	4701	cccccitggcttcagtcatcccaacgacaagatgtatggtaatgtat
genomic_hu	1	-----
cDNA_rat	4751	agtgcagttgttatcactcaacaacotctttctgccataactgaact
genomic_hu	1	-----
cDNA_rat	4801	gttgagaagtacacccagactttggaaatacaacagctttggaaacaa
genomic_hu	1	-----
cDNA_rat	4851	cgttgtgagcaaattcacaggagagtaccacacgtgaaaagagcctcagac
genomic_hu	1	-----
cDNA_rat	4901	acaccaccaccacttcgtcagcagtggggcgcccccagtgcacactccttc
genomic_hu	1	-----
cDNA_rat	4951	cccaccccttttactaagggtgtggttacagacagcaagtccatcatcg
genomic_hu	1	-----
cDNA_rat	5001	ctttccagatgacgtcaaataaggtggtcaccatataatgaatcttcaagg
genomic_hu	1	-----
cDNA_rat	5051	cacaatacagatctgcagcaaccctcagcagaggctagccccatctgt
genomic_hu	1	-----
cDNA_rat	5101	gatcataactgaaaccactgacttcccttaatctgtttccatccactt
genomic_hu	1	-----
cDNA_rat	5151	ctgtgccagcactaaggtagataaaccacagaattctaaatggaaagcc
genomic_hu	1	-----
cDNA_rat	5201	tctccctggccagaacacaaatatcagtcagtcataactccgaaaccat
genomic_hu	1	-----

cdNA_rat	5251	tgagaaggcaaaaggccaggcagtaagcatgtccccccacctcagcctc
genomic_hu	1	-----
cdNA_rat	5301	cagaggccagcactcatgcctcacactggataacacagaagcatgcagaa
genomic_hu	1	-----
cdNA_rat	5351	aagagtgttttataagaacctggtaaaaacocaaacttccaaacatct
genomic_hu	1	-----
cdNA_rat	5401	gccttacgtctctactaagacttotattgaaaaagccaagaataattg
genomic_hu	1	-----
cdNA_rat	5451	gaggaaggotgcaagcttacagttccagcttaattcagaogttttctt
genomic_hu	1	-----
cdNA_rat	5501	ccttgcaggotgttggagaccactgcccattatccactggaccagagt
genomic_hu	1	-----
cdNA_rat	5551	ttcatcaggantgaaatatccaaaggacacagaaaagccggttccacg
genomic_hu	1	-----
cdNA_rat	5601	tgottccaaatggcaccttgcattccagagggcagtttccaggaccgt
genomic_hu	1	-----
cdNA_rat	5651	ggacagtacctgtctgcatttaatccactggcgttagaccatttca
genomic_hu	1	-----
cdNA_rat	5701	tgtottttgtctgtgttttacccggcaaggattttggacagacatg
genomic_hu	1	-----
cdNA_rat	5751	tcaaggagatcacagttacatttggaaagtactgtggaaactsaagtgcaga
genomic_hu	1	-----
cdNA_rat	5801	gtggagggtatgccgaggcctacggttctggatacttgcaaaaccaac
genomic_hu	1	-----
cdNA_rat	5851	ggtggtctcagaaacggcaaggaaagcagaaaggctggtaacacctg
genomic_hu	1	-----
cdNA_rat	5901	atggacattgatcatataatctgagtctttatgatcgtggttttac
genomic_hu	1	-----
cdNA_rat	5951	asgtgtgtggccagcaacccatctggccaggattcaactgttgtaagat
genomic_hu	1	-----
cdNA_rat	6001	acaagtcatcacagctccctgtcattatacgtcaaaaaggcaaggca
genomic_hu	1	-----
cdNA_rat	6051	tctgtgggttttagtgtggaaactgcccactgcactgcacaaaggaa
genomic_hu	1	-----
cdNA_rat	6101	actccccagcttagtgttcaactgggtctttatgatggactgaactaaa
genomic_hu	1	-----
cdNA_rat	6151	accattgcaggtagtcattccagattttctgtatccaaatggaaactc
genomic_hu	1	-----
cdNA_rat	6201	tgtatataagaaggcatacgtcccttcagtggggcacttatgagtgcatt
genomic_hu	1	-----
cdNA_rat	6251	gccaccagctcctcaggctcagagagaaggtagtgattcttactgtgga
genomic_hu	1	-----

cDNA_rat	6301	agagggagagacaatccocaggatagaaactgccttcagaaatggactg
genomic_hu	1	-----
cDNA_rat	6351	aggtaatttgggtgagaaattactactgaactgctcagctactgggat
genomic_hu	1	-----
cDNA_rat	6401	ccaaagccatagaataatctggaggctgccatccaggctgtcatogacca
genomic_hu	1	-----
cDNA_rat	6451	gtggcacagaaatGGGCAGCCGAATCCACGCTCACCAAATGGATCCTGG
genomic_hu	1	-----TGGCAGCTGGATCCACGCTCACCTAATGGATCCCTGt
cDNA_rat	6501	tggttgggTCAGTGACGGAAAAAGACGCTGGTGA
genomic_hu	40	ttattggatCAGTAACAGAAAARGACAGTGGTGTCACTTGTGTGGCA
cDNA_rat	6551	AGAAACAAAATGGGAGATGACCTAGTCCTGATGCATGTccgcctgAGATT
genomic_hu	90	AGAAACAAAATGGGGAGATGATCTGATACTGATGCATGTTagcctaAGACT
cDNA_rat	6601	GACACCTGCCAAAATTGAACAGAACAGTATTTAagAAGCAAGTGCCTCC
genomic_hu	140	GAAACCTGCCAAAATTGACCACAAAGCAGTATTTAgaAAGCAAGTGCCTCC
cDNA_rat	6651	ATGGGAAAGATTTCAAGTGTGACTGCAaggcctctGGCTCCCCGTGCCT
genomic_hu	190	ATGGGAAAGATTTCAAGTAGATTGCAagcttccGGCTCCCCAGTGCCTA
cDNA_rat	6701	GAGGTATCCTGGAGTTGCCTGATGgacagtgcTCAACAATGTAGCCCA
genomic_hu	240	GAGATATCTGGAGTTGCCTGATGgaaaccatgtCAACAATGCAATGCA
cDNA_rat	6751	AGCTGATGACAGTGGCTATAGGACcaagaggtaACCCCTTTCAACATG
genomic_hu	290	AGCCGATGACAGTGGCCACAGGACTaggagatataACCCCTTTCAACATG
cDNA_rat	6801	GAACCTTGTATTCACAAACGTTGggatggcaGAGGAAGGGAGATTATATC
genomic_hu	340	GAACTTATACTTCAACAAAGTTGggtagcgGAGGAAGGGAGATTATACT
cDNA_rat	6851	TGCTCTGCCAGAACACCTTAAGGAAAGATGAGATGAAAGTCCACCTAAC
genomic_hu	390	TGCTATGCCAGAACACCCCTAAGGAAAGATGAAATGAAGGTCCACTTAAC
cDNA_rat	6901	AGTTCTAACAGCcatcCCACGGATAAGGCAAAGCTACAAAGACCCACATGA
genomic_hu	440	AGTTATAACAGCgtctCCCGGATAAGGCAAGTAACAAACCAACAAGA
cDNA_rat	6951	GGCTCAggGCTGGAGAACACGCTGTCCTTGACTGCGAGGTCACTGGGaa
genomic_hu	490	GAATCAaaGCTGGAGACACAGCTGTCCTTGACTGTGAGGTCACTGGGat
cDNA_rat	7001	ccaaagccaaatgtATTTTGGTTGCTGCCCTCCAAatgtcATTTCATT
genomic_hu	540	ccaaacccaaaaaATTTTGGTTGCTGCCCTCCAAatgtcATTTCATT
cDNA_rat	7051	CTCCAATGACAGGTTCACATTCAATGCCAATgaaaCTTGTCATCCATA
genomic_hu	590	CTCCATTGATAGGTACATTCAATGCCAATgggtCTTGACCATCAACA
cDNA_rat	7101	AAGTGAAACCACTTGACTCTGGGGActatgtgtcgtagtcagAAATCCT
genomic_hu	640	AAGTGAAACTGCTCGATTCTGGAGAgtaactatgtgttagccccgaAAATCCC
cDNA_rat	7151	AGTGGGGATGACACTAACAGACATACAAACTGGAcattGTCTCTAAACCTCC
genomic_hu	690	AGTGGGGATGACACCAAAATGTACAAACTGGAtgtgGTCTCTAAACCTCC
cDNA_rat	7201	ATTAATCAATGCCCTGTATGCAAACAAagACTGTTATTAAGCCACAGCCa
genomic_hu	740	ATTAATCAATGGTCTGTATACAAACAgaaACTGTTATTAAGCCACAGCtg
cDNA_rat	7251	ttcgccacTCCAAAAAAATACTTGTACTGCAAGAGCAGATGGATCCCATct
genomic_hu	790	tgagacatTCCAAAAAAACATTGACTGCAAGAGCTGAAGGGACACCATct
cDNA_rat	7301	tcccaaggcacGTTGATTATGCCAGGCAATATTTCTCCAGCTCCATA
genomic_hu	840	cctgaagtcatGTGGATCATGCCAGACAAATTTCTCACAGCCCCATA

cDNA_rat	7351	CttTGGAAAGCAGAGTCACGGTCCATCCAAATGGAAACCTTGGAgatgagga
genomic_hu	890	CtATGGAAGCAGAAATCACAGTCATAAAAATGGAAACCTTGGAAatttagga
cDNA_rat	7401	acatccGGCTTCTGACTCTGGCGACTTCACCTGTGTGGttcgagcgag
genomic_hu	940	atgtgaGGCTTCAAGATTCAGCCGACTTTATCTGTGTGGcccgaaastgaa
cDNA_rat	7451	ggaGGAGAGAGTGTGTTAGTAGCAGTTAGAAGTCCTAGAAATGCTGAG
genomic_hu	990	ggtGGAGAGAGCGTGTGTTAGTACAGTTAGAAGTACTGGAARTGCTGAG
cDNA_rat	7501	AAGACCAACATTCAAGAAACCCATTCAACGAAAagtcatcgcccaagctg
genomic_hu	1040	AAGACCGACATTAGAAATCCATTAAATGAAAAatagtgtgccagctgg
cDNA_rat	7551	gcaagccccgtAGCACTGAJCTGCTCTGTGGATGGAAACCCCCCACCTGAA
genomic_hu	1090	gaaagtccacAGCATTGAATTGCTCTGTGATGGTAACCCACCCACCTGAA
cDNA_rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTGCTAACAGACCCAcacaa
genomic_hu	1140	ataatCTGGATTTACCAATGGCACACGATTTCCAATGGACCAacaaag
cDNA_rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTACAAAagcaa
genomic_hu	1190	ttatcaGTATCTGATAGCAAGCAATGGTCTTTATCATTTCTAAAacaa
cDNA_rat	7701	ctcggaacaagtCAGGGAACTATCGCTGTGCAGCCAGGAATAAGGTTGGC
genomic_hu	1240	ctcgggaggatgCAGGAAAATATCGCTGTGCAGCTAGGAATAAGGTTGGC
cDNA_rat	7751	TAcatcgagaaaactcatcctgttagagATTGGGCAGAACGCCAGTCATTCT
genomic_hu	1290	TAtattgagaaaatttagtcatattagaaATTGGCCAGAACGCCAGTTATTCT
cDNA_rat	7801	GACATACGAACCAGGGAtggtaagagqcgTCAGTGGGAACCGTTATCAC
genomic_hu	1340	TACCTATGCACCAGGGAcagtaaaaggoTCAGTGGAGATCTCTATCAC
cDNA_rat	7851	TGCATTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
genomic_hu	1390	TGCATTGTGTCTGATGGaATCCCTAACGCAAATATCAAATGGACTATG
cDNA_rat	7901	CCGGGTGGccATGTAATCGACAGGCCCTCAAGTGGATGGAAAATACATAct
genomic_hu	1440	CCAAGTGGttATGTAGTAGACAGGCCCTCAAATTAAATGGAAATACATATT
cDNA_rat	7951	GCATGAAAATGGCACGCTGGCATCAAAGCAACACAGCtcacgaccaAG
genomic_hu	1490	GCATGACAATGGCACCTTAAGTCAATTAAAGAAGCAACAGCttatgacagAG
cDNA_rat	8001	CAAATTATATCTGTAGGGCTCAAAACAGTGTGTTGGCAggcagttattagc
genomic_hu	1540	AAAATATATCTGTAAAGCTCAAAATAGTGTGTTGGTCAtacactgattact
cDNA_rat	8051	gtgtCAGTGTGGTTGTGGCTTACCCCTCCCCAATCATAAAActacctACC
genomic_hu	1590	gttcCAGTAATGTGTTAGCCTACCCCTCCCCAATTACAAAAtogtccACC
cDNA_rat	8101	CAGGAACATGTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
genomic_hu	1640	CAGGAGTATTGTCACCAGGACACGGGCAGCCTttCAGCTCCACTGTGTGG
cDNA_rat	8151	CCTTGGGAATCCCCAAGCCaaAAGTCACCTGGAGACGCCAAGACACTCC
genomic_hu	1690	CCTTGGGAGTTCCAAGCCAgAAATCACATGGGAGATGCCGTGACCAACTCC
cDNA_rat	8201	CTGCTCTCAAagcaacagcaagaasaccCATAGAAGTGGAGATGCTTCA
genomic_hu	1740	CTTCTCTCAccggcaagttaaagagaggacacATGCCAGTGGAGCAGCTTCA
cDNA_rat	8251	CCCACAAGGTACGCTgGTCATTCAAGAATCTCAAACCTCGGATTCCGAG
genomic_hu	1790	CTTACAAGGTACCTaGTCATTCAAGAATCCCCAACCTCCGATTCTGGga
cDNA_rat	8301	tcTATAAGTGCAGAGCTCAGAACCTACTTGGactTGATTACGCAACAACT
genomic_hu	1840	taTACAAATGCACAGCAAAGAACCCACTTGGtagTGATTATGCAGCAACG
cDNA_rat	8351	TACATCCAGGTACTCTGACAGGAAGgggggagactaaaattcaacagaagt
genomic_hu	1890	TATATCAAGTAATCTGACATGAAataataaagtcaacaaacatctgggca

cDNA_rat	8401	ccacatocacagggtTTTATTTGGAAAGAAGTTAACAAAGGCAGCCA
genomic_hu	1940	gaa-----TTTATTTGGAAAGAAGTTAACAAAGGCAGCCA
cDNA_rat	8451	TAGGCATGTAAATGAGtcTGAATACTTACAGTATTAAATTACAATGG
genomic_hu	1979	TAGGCATGTAAATGAattTGAATACTTACAGTATTAAATTACAATGA
cDNA_rat	8501	ACATGCgtga-----GACTTGTAAATGAAAGCATTGTGAACGTGAAaccg
genomic_hu	2029	ACATGCaaaataaaagGACTTGTAAATAAATGCATTATGAACGTGAtgata
cDNA_rat	8546	agtctctg-----TGGATCTCAAAGCAAACTCTTAACCTTAAGGCACTTg
genomic_hu	2079	otgattttatTTAATGGATCTCAAACAAACTTTAACCTTAAGGCACTTt
cDNA_rat	8591	ATTTTGCCAACAAATAATAACAAACAttaagagaaaaaaatgtcCACTA
genomic_hu	2129	ATTTTGCCAACAAATAACAATAAAACAAacattgaaacggtt---CACTA
cDNA_rat	8641	CGAAATAACAAACGGCTAATGCACCTGAATTctcagtaaaaagacottc
genomic_hu	2175	TAAAATAACAAATGGCTAATGTACCTGAATTttcagtaaaaaaaatgaac
cDNA_rat	8691	tctcgctaacagttgCCAGCTGCCTCGTGTGTTCCCTACCAATGTCAC
genomic_hu	2225	ttctataa-----CCAGTTGCCTAGTGTCCACCTCTATCAATGTAC
cDNA_rat	8741	AAACAtogcacacagggtGAATGGAGTCACGGGAAAGATTAAGTTGCG
genomic_hu	2268	AAGCAAtggcactca---GAACAGAGACAATGGAAAAATATTAAATCTGCA
cDNA_rat	8791	GTctgtgtatctca-----ATGTACAAATATTCTGtcncTG
genomic_hu	2314	ATCtttatgtaaatttaccatcctgATGTATAAAATATTTG---TG

exon1 (2342-2397)

cDNA_rat	8791	GTctgtgtaaatctca-----ATGTACAAATATTCTGtcncTG
genomic_hu	2314	ATCtttatgtgtaaatttaccatcctgATGTATAAAATATTTG---TG
cDNA_rat	8829	GTTTATAAACATTGATAAAACGGAAAAAAAAAAAAAAaaaaaaaaaaaaa
genomic_hu	2360	GTTTATAAACATTGCTAACCTACAGAAAATAAGcactgaactgtc

(Genomic human OCP: SEQ ID NO:4)
(cDNA rat: SEQ ID NO:5)

Figure 11

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length	Remarks
1	1	208	208	69	No valid splice site found upstream this exon
2	277	429	153	18	
3	447	485	39	1561	
4	2046	2244	199	1351	
5	3595	3724	130	3254	
6	6978	7359	382	4123	
7	11482	14903	3422	38	
8	14941	15307	367	51	
9	15358	15825	468	1039	
10	16864	~17760	~897	?	Last exon might be complete
11	?(1)	2317	~231 7	25	Exon is not complete and start site is not known
12	2342	2397	56	-	

FIGURE 12

608 Human translated nucleotide sequence (ORF)

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTGGCTCCTTGTG
 GATCTGCCTGGTCGCCACCCCTGGGGCAAGGCCTGCCTGCCGCTGTG
 CCTGTTATATGCCTACGGAGGTACACTGCACATTCGGTACCTGACTTCCA
 TCCCAGACAGCATCCGCCAATGTGGAACGCATCAATTAGGATACAAC
 AGCTTGGTAGATTGATGGAAACAGATTTCTGGCTGACCAAACGTGGA
 GTACTCATGCTCACAGCAATGGATTACACAAATCCCTGACAAGACCTT
 CTCAGATTGCAAGGCCTTGCAAGGTTAAAAATGAGCTATAATAAGTCC
 GAAAACCTCAGAAAGATACTTTATGGCTCAGGAGCTTGACACGATTG
 CACATGGACACAAACAAATTGAGTTATAAAACCCAGAGGTTTTATGG
 GCTCAACTTCTCCGCCGGTGCACITGGAAAGGAAATCAGCTCACTAAGCT
 CCACCCAGATACATTGCTCTTGAGCTACCTCCAGATATTAAAATCTC
 TTGATTAAGTTCCTATACCTGCTGATAACTTCCTGACCTCCCTCCCTCAA
 GAGATGGTCTCCTATATGCTGACCTAGACAGCCTTACCTGATGGAAA
 CCCATGGACCTGTGATTGCCATTAAAGTGGTTGTCTGACTGGATAACAGCC
 AGATGTAATAAAATGCAAAAAAGATAGAAGTCCCCTAGTGCTCAGCAGT
 GTCCACTTGCATGAACCTAGGACTCTAAAGGCAGCCGTTAGCTATG
 GTCTCAGCTGCACTTCCAGTGTGCCAAGCCAACCATTGACTCATCCCTG
 AAATCAAAGAGCCTGACTATTCTGAAAGACAGTAGTTCTGCTTCATCTC
 CCCCAGGTTCATGGCACCTTGGCTCCCTACTTGAATATGACAGAT
 CAGTCTGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAG
 GACATCACCATTGCATTCACTGAAGAAAATGACTACATCGTGTAAATA
 CTTCATTTCAACATTTGGTGTGCAACATAGATTACGGTCACATTCAAGC
 CAGTGTGCAAATTGGCTTGTACAGTGATCTCCTGATACTAGAAA
 GGAGCCACTTGTAGTGAAGACATTACCAACATAGAGGCAGATCTCAG
 GTGGCTCTAACGCCTGAAGACATTACCAACATAGAGGCAGATCTCAG
 AGCAGATCCCTTGGTTAATGCAAGACAAATTCCCTGAGCTGAACA
 GAACTGCCACCACATTCACTGACATTACAGATCCAGTACTCCAGTGTGCT
 CAAATCACTTACCAAGAGCAGAGATGAGGCCAGTGAACACAAATGGA
 CTATGATTCAGGGATAACAATACTAAGCTGGAACATACTGTCTGGTA
 GGTGGAACCGTGGCTGAACGCCAGGCAAGGAGACCCACCCACA
 CGTGGATTGGCTCTAGTGTGAAAGTGAAGGAGAGCCCTATGTCA
 GTGAGGATGGACGGATCTTAATAGACAAAAGTGGAAAATGGAACTCCA
 GATGGCTGATAGTTGACACAGCGTATATCACTGTATAAGCAGCAATT
 ATGATGATGCAGATATTCTCACCTATAGGATACTGTGGTAGAACCTTGG
 TCGAAGCCTATCAGGAAAATGGGATTCTACACAGTTTCTATTGGTGAA
 AACCTGATCTCCATGCCATTCTACTGGTATCCAGATGCCCTATTAGC
 TGGGTTATCCAGGAAACAACTGCTCTACAGTCATCAAGAGACAAGAA
 AGTCTAAACAATGGCACATTAAGAATATTACAGGTACCCGAAAGACC
 AAGTTTATTATGCTGTGGCAGCCAACCCATCAGGGGTTGATTTTGAA
 TTTCCAAGTTCACTGCAAGATGAAAGGACAAGGCCCTGGAGCATGAT
 GGAGAAACAGAGGGATCTGGACTTGTGAGTCAATCCTATTGCTCATCT
 TAAGGAGGCCACAGGTGACAACCTCGTACATCTGCTGTGAGGAGCTG
 AGGTGGAAAACACACCTCAAGCACAAGTAAGAGGGACAACATCGGGA
 ATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTAGGGAGA
 ATAGGAGGCATTGCCCTCTGCTAGGAGAATTGACCCACAACATTGG
 GCGGCACTGTTGGAGAAAGCTAAAAGAATGCTATGCCAGACAAGCGAG
 AAAATACCAAGTGAGGCCACCCCAAGTGGTACCCAACTCCAAACATA
 CCTGGTGAAGAAGACGATTCCCTAGGCATGCTCGCTACATGAGGAATT

TATGGTCCCGGCCACTAAAGCTTGAACCTTCCAGCAAGGACAGTGACTG
 CTGACTCCAGAACAAATATCTGATAGTCTATGACAAACATAAATTATGGC
 ACAGAATTCTCTCTGTTGTGAATTCAAAATACTACCCACCTGAAGAACCC
 ACAGATTCAAACATGTCTACTGCTATTAAAACATACAGCCATGTCAAAGAA
 TATAAACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATT
 CATCCACTGTCTTCCACTGCTACTGGAGCAACTGAATTCAAGGACTCTG
 ACCAGATGGGAAGAGGAAGAGAGCATTCCAAAGTAGACCCCCAATAAC
 AGTAAGGACTATGATCAAAGATGTCATGTCAAATGCTTAGTAGCACCA
 CCAACAAACTATTAGAGTCAGTAATACCACAAATAGTCATCACAGACA
 TCTGTAAGAGAAGTGAAGTGAACCCAGGCACAATCACTTCTATTCTCACAC
 TACTCAAATACCTAGCACCTCCACGTTCCCTCAGATCCACACACAGCTGC
 TCATTCTCAGTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGT
 CAGACGCTTGGAGGCAGAGGAAAATTGGCGAAGGGGGCGGATTATC
 AGCCCCATATAGAACTCCAGTCTCGCAGGGCATAGATACAGCATTTTCAG
 GTCAACAAACAGAGGGTCTTCTGAAAAAAAGCACTACTGCAATTCTCAGCCA
 CAGTGCATGACATGTGACATGTCTGCTGTCTCCAGGGAGAGGCTCACCA
 CTGCCACAGCAGCATGTCATTCAGTCAAGTGCCTCCATCACCTCCCCA
 AAGCTGACATIGCTAGAGTCCCATCAGAAGAGTCTACAACCTCTAGTCCAG
 AATCCCACTATTACTACTTGGAAACAAACCCAGTGTAGAGAAAACACACC
 CACAATAAAATATTCAAGGACTGAAATTCCAAGTGACTCCAACGGTGTG
 CAGTCATGACATATGCTCAACATCCATACCCATGGAAAAAAACTCACAAA
 GTAAACGCCAGTACCCACGTGTCTAGCACCAATGAAGCTAAAGAGA
 TTCACTGATTACATCGTCACTTCAAGGTCTAGCACCAAGCCACCAATGAC
 TATTATAGCCATTACAAGGTTTCAGAAGGAAAATTCCCTGGCAACAGA
 ACTTTGTAATAACCATAACCCAAAAGGCAGATTAGGAATCAACATAAA
 GTTAGTTACAAAAAGCACAGCTGTGATGCTTCTAAACATCTCTGCT
 TTACCCACAGACAAAGTTCCCTTCCATTCAACACTTCAACAAAGT
 GTGATGCAAATCCATTAACCTTGACTACCGCTCACACACTACGACC
 AAAACACACAATCTGGAAGTCTCCAACAAAGAAGGAGCTTCCCTTCCC
 ACCCTTAACCCATGCTTCTAGTATTATAAGCAAAGACTCAAGTACAA
 AAAGCATCATATCAACGCAAACAGCAATACCAAGCAACAACCTCCTACCTC
 CCTGCATCTGTCACTTATGAAACCCAAACAGAGAGATCTAGAGCACA
 AACAAATACAAGAGAACAGGAGCCTCAAAAGAAGAACAGGACTGACCCA
 AACATCTCCAGACCAGAGTCTGGCTTCACTACACCCACTGCTATGACA
 CCTCTGCTCTGGCATTCACTCATCCCCACCAAGAAAACAACACTGGGATT
 TCAAGCACAATCAGTTTCATTCAAGAACACTTAACTGACAGATGTGATT
 GAAGAACTAGCCAAAGCAAGTACTCAGACTTGAAGAGGACAATTGCTTC
 TGAAACAATTGTCAGAACATCACCCAGAGTACCTGACTACCCACTGCTATGACA
 CATCATTAGACACTCCCATACCCATTCTGAGCAGCAGTGTCTACTCTAA
 TGCCAGTTCCCATCTCCCTCCCTTACTCAGAGAGCAGTTACTGACACAC
 GTGGCGACTCCCATTCCGGCTTATGACAAATACAGTGGTCAAGCTGCAC
 GAATCCTCAAGGCACAATCTCAAATGCCAGTTCAAAATTGGAAACCACT
 CACTTCATCTACCTCTAACTGTTACATTCTACTCCCATGCCAGCACTAAC
 AACAGTTAAATCACAGAACCTCAAATTAACTCCATCTCCCTGGCAGAAT
 ACCAATTGGCACAAACCAACTCAGACATTGCTAAAAAGGCAAAAG
 CCAGAAGTAAGCATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCT
 TGTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGAGTGACTIONGATA
 AGAAACCAGTCAAGAAGCAACAACTTCAAACCTCCCTCCCTTGACTCTT
 TGTCTAGGTATAATTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCA
 AGTTTACTATTCCAGCTAACTCAGATGCCCTTCTCCCTGTGAAGCTGTT

GGAAATCCCTGCCACCATTGACCAGAGTTCAAGGACTTGTATTA
 TCTAGAGGAACCAGAACATAGCAGGGTCCAGGTCTCCCAATGGTACCC
 GTCCATCCAGAGGTGAAATTCAAGGACCGCCGACAGTACTTGTGTTCCG
 CATCCAATCTGTTGGCACAGACCACCTCATGTCACCTTGTCTGTT
 CCTATCCTCCAGGATCTGGAGAGACGTACCAAAGAGATCACAGTCAT
 TCCGGAAGCACTGTGGAACGTGAAGTGCAGAGCAGAAGGTAGGCCAAGCC
 CTACAGTTACCTGGATTCTGAAACCAAACAGTTGTCTCAGAATCATCCC
 AGGGAAAGTAGGCAGGCTGTGGTACGGTACGGAACATTGGCTCCAC
 AATCTCAGTATTATGACCGTGGTTTACAAATGTGTGGCAGCAACCCA
 GGTGGCCAGGATTCACTGCTGGTAAATACAAGTCATTGAGCACCACC
 TGTATTCTAGAGCAAAGGAGGCAAGTCATTGAGGCACTGGGGTGA
 GTTAAAATGCCCCTGTACTGCAAAAGGAACCTCTCAGCCCAGCGTTACT
 GGGTCTCTGATGCCACTGAAGTGAACCTACAGTTACCAATTCCA
 AGTTGTTCTTATTCAATGGGACTTGTATATAAGAAAACCTAGCCTTT
 CAGACAGGGCACTTATGAATGCATTGCTACCAGTCCACTGGTCCGGAG
 CGAAGAGTAGTAAATGCTTACAATGGAAGAGCGAGTGAACAGCCCCAGGA
 TAGAAGCTGCATCCCAGAAAAGGACTGAAGTGAATTGGGACAAATTA
 CTACTGAACTGCTCAGCCACTGGGAGCCAAACCCAAATAATGTGGAG
 GTTACCATCCAAGGCTGTGGTCGACCAGTGGAGCTGGATCACGTCTACC
 CTAATGGATCCCTGTTATTGGATCAGTAACAGAAAAGACAGTGGTGT
 TACTTGTGTGGCAAGAAAACAAATGGGGATGATCTGATACTGATGCA
 TGTAGCCTAAGACTGAAACCTGCCAAATTGACCAAGCAGTATT
 GAAAGCAAGTGCCTCATGGGAAAGATTCCAAGTAGATTGCAAGCTTCC
 GGCTCCCGAGTGCAGAGATATCTGGAGTTGCTGATGGAACCATGAT
 CAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGAT
 ACCCTTTCAACATGGAACCTTATACTTCAACAAAGTTGGGTAGCGGA
 GGAAGGAGATTATACTGCTATGCCAGAACACCCCTAGGGAAAGATGAA
 ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCGATAAGGCAGAG
 TAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAGCTGCTTGC
 GTGAGGTCACTGGGATCCAAACCAAAATATTGGTGTGCTGCCTTCC
 AATGACATGATTCTCTCCATTGATAGGTACACATTCTGCAATGGG
 TCTTGACCATCAACAAAGTGAACACTGCTGATCTGGAGAGTACGTATG
 TGTAGCCGAAATCCCAGTGGGATGACACCAAAATGTACAAACTGGATG
 TGGTCTCTAAACCTCATTAAATCAATGGTGTGTATACAAACAGAACTGTT
 TAAAGGCCACAGCTGTGAGACATTCCAAAAACACTTGTACTGCAGAGCT
 GAAGGGACACCATCTCTGAAGTCATGTGGATCATGCCAGACAATT
 CCTCACAGCCCCATACTATGGAAGCAGAACATCACAGTCATAAAATGGAA
 CCTTGGAAATTAGGAATGTGAGGCTTCAAGATTGCGACCTTATCTGTG
 TGGCCGAAATGAAGGTGGAGAGAGCGTGTGTAGTACAGTTAGAAGTA
 CTGGAAATGCTGAGAAGACCGACATTAGAAATCCATTAAATGAAAAAAT
 AGTTGCCAGCTGGAAAGTCCACAGCATTGAATGCTGTGATGGTA
 ACCACCAACCTGAAATAATCTGGATTACCAATGGCACACGATT
 ATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTCTTATCA
 TTCTAAAACAACCTGGGAGGATGCGAGAAAATATCGCTGTGAGCTAGG
 AATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATTGGCAGAA
 GCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAGGCATCAGTGGAG
 AATCTCTATCACTGCATTGTGTCTGATGGAATCCCTAACGCCAATATCA
 AATGGACTATGCCAAGTGTATTGTAGTAGACAGGCCAACATTAAATGGG
 AAATACATATTGCATGACAATGGCACCTAGTCATTAAGAAGCAACAGC
 TTATGACAGAGGAAACTATATCTGTAAGGCTAAATAGTGTGGTCATA

CACTGATTACTGTTCCAGTAATGATIGTAGCCTACCCCTCCCCGAATTACAA
ATCGTCCACCCAGGAGTATTGTCACCAGGACAGGGGCAGCCTTCAGCTC
CACTGTGTGGCCTTGGGAGTTCCAAGCCAGAAATCACATGGGAGATGCC
TGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGAAGTG
AGCAGCTTCACTTACAAGGTACCTAGTCATTAGAATCCCCAACCTCC
GATTCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGTAGTGATTA
TGCAGCAACGTATATTCAAGTAATCTGA

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

Figure 14

rat_cDNA	AGCCTTACTAGATTGACAGAAAACGACTTGTATGGCTGAGCAAACCTGGAGTTACTCATG
human_5+3_corrected	AGCTTGGTAGATTGAGTGGAAAACGATTTCTGGCTGACCAAACCTGGAGTTACTCATG
mus_cDNA_5	AGCCTCACTAGATGACAGAAAATGACTTTCTGGCTGAGCAAACCTGGAGTTACTCATG
*****	*****
rat_cDNA	CTGCACAGTAATGGCATTCACAGAGTCAGTGACAAGAAGCTCTCGGGCTTGCAGFCCTTG
human_5+3_corrected	CTTCRACAGCAATGGCATTCACRCAATCCTGACRAGACCTCTCAGATTTGCAGGCTTG
mus_cDNA_5	CTGCACAGCAATGGCATTCACAGAGTCAGTGACAAGAAGCTCTCGGGCTTGCAGFCCTTG
*****	*****
rat_cDNA	CAGGTCTTAAAATGAGCTATAACAAAGTCCAATCATTOGAAGGATACTTTTACGGAA
human_5+3_corrected	CAGGTCTTAAAATGAGCTATAATANAGTCCAAAACCTTCAGAAAGATACTTTTATGGC
mus_cDNA_5	CAGGTCTTAAAATGAGCTATAACAAAGTCCAATAATGAGAAGGATACTTTTATGGAA
*****	*****
rat_cDNA	CTCGGGAGCTTGGTCGGTGCACCTGGATCACAAACACATTGAAATTCAACCCGTGAG
human_5+3_corrected	CTCAGGAGCTTGAACAGATTGCACCTGGACCAACAACATAATTGAGTTTATAACCCAGAG
mus_cDNA_5	CTCAGGAGCTTGAACCGGTTGCACCTGGATCACAAACACATTGAGTTTATAACCCGGAG
*****	*****
rat_cDNA	GCCTTTTATGGACTTACCTCGCTCGCTGGTACATTAGAAGGAAACCGGCTCACAAAG
human_5+3_corrected	GTTCCTTTATGGGCTCAACITTCCTCGCTGGTGCACITTGAGGAAATCAGCTCACTAAG
mus_cDNA_5	GCCTTTTACGGACTCACCTGCTCCGCTGGTACATCTAGAAGGAAACCGGCTCACAAAG
*****	*****
rat_cDNA	CTCCATCGACACATTTGTCATTAAGCTATCTCAGATAATTAAACCTCTTCATT
human_5+3_corrected	CTCCACCCAGATACTTTGTCCTTGGAGCTACCTCCAGATAATTAAACCTCTTCATT
mus_cDNA_5	CTCCATCGACACATTTGTCCTTGGAGCTATCTCAGATAATTAAACCTCTTCATT
*****	*****
rat_cDNA	AAGTACCTGTTCTGCTGATAACTTCCT-GACCTCCTCCAAAAGAAATGGTCCTCTA
human_5+3_corrected	AAGTTCCTATACCTGCTGATAACTTCCT-GACCTCCTCCCAAAAGAGATGGTCCTCTA
mus_cDNA_5	AAGNACCTGTAATTGATGATAACTTCATTGACCTCCTCCAAAAGAAAATGGTCCTCTC
***	***
rat_cDNA	CATGOCAAAACCTAGAAAGCCTGTATTGAGTGGAAACCCATGGACCTGTGACTGCCATT
human_5+3_corrected	TATGCTCTGACCTAGACAGCCCTTACCTGCTATGGAAACCCATGGACCTGTGACTGCCATT
mus_cDNA_5	TATGOCAAAACCTAGAAAGCCTTACCTGCTATGGAAACCCATGGACCTGTGACTGCCATT
*****	*****
rat_cDNA	AAAGTGGTTGCTGAGTGGATGCCAGGGAAACCCAGATAATAATAAAATGCRAGAAAGACAG
human_5+3_corrected	AAAGTGGTTGCTGACTGGATACAGGNNNNNNNOCAGATGATAATAAAATGCAAAAAGATAG
mus_cDNA_5	AAAGTGGTTGCTGGAGTGGATGCCAGGGAAACCCAGGTA-ACTAATCTGT-----TTGTTG
*****	*****
rat_cDNA	AAGCTCTTCCTGCTCAGCATGCTGGCTCCCT-TTGCATGAAACCCAGGATCTCTAAAGCA
human_5+3_corrected	AAGTCCTCTAGTGCTCAGCATGCTGGCTCCCT-TTGCATGAAACCCAGGATCTCTAAAGCA
mus_cDNA_5	TTTCCTTTTATARKAGTTTCTCAATTCTATTAGAATGATACTCCAAAAGTC-
*****	*****
rat_cDNA	GACCTTTGCTATGGTACCATCTGGACCTTCTATGTCATAAAAGCCAAACCTTGATCCAT
human_5+3_corrected	AGCCCTTACGCTATGGTCTCAGCTGCACCTTCTCAGTGCTGCCAAAGCCAAACCTTGATCCAT
mus_cDNA_5	-CCCGATAACCTCCCCCCC-----CTTCCCTACCTACCCATTC-CCATTTCGGC
***	***
rat_cDNA	CACTGAAGTCATAAGAGCCTGGTACTCAGGAGGACAAATGGATCTGCTTCCACCTCCTC
human_5+3_corrected	CCCTGAAATTCATAAGAGCCTGACTATTCCTGAGACAGTACTCTGCTTCTATCTCTCCC
mus_cDNA_5	CCTGGCATTCCTC-----
*****	*****
rat_cDNA	AAGATTTCATAGAACCCCTTGGCTCTTGTCTTGAACTGACANANNTCTGGAAATA
human_5+3_corrected	AAGGTTTCATGGCACCCCTTGGCTCCCTACTTTGAATATGACAGATCAGTCAGTCAGTC
mus_cDNA_5	-----
rat_cDNA	AGGCCGACATGGTCTGACTATCCTAAAGCCATCAAGGACATCACCAACTGCATTCACTG
human_5+3_corrected	AAGCTAACATGGTCTGACTATCCTAAAGCCCTCAAGGACATCACCAACATTGCAATTCACTG
mus_cDNA_5	-----

rat_cdna human_5+3_corrected mus_cdna_5	AAGAAAATGACTACATCATGCTAAATGGTCATTTCCACAAATCTTGCTGCAAGTGAG AAGAAAATGACTACATCGTCTAAATCTCATTTCAACATTTGGCTGCAACATAG
rat_cdna human_5+3_corrected mus_cdna_5	ATTATAATCACATCCAGCCAGTGTGGCAACTCTGGCCTTATACAGTGCACITCTCTGA ATTACGGTCACATTCAGCCAGTGTGGCAATTTGGCTTGTACAGTGAATCTCTCTGA
rat_cdna human_5+3_corrected mus_cdna_5	TACTAGAAAGGAAGCCCAGCTTACCGAGACTCCTTCACITGCTCTAGATATAACAGG TACTAGAAAGGAGCCACTTGTCTTGTAAACACCGCAGCTCTATTACAATATAACAGG
rat_cdna human_5+3_corrected mus_cdna_5	TGGCTCTTAGGCTGAAGACATTTTACCACTAGAGGCTGATGTCAGAGCAGACCTT TGGCTCTTAAGGCTGAAGACATTTTACCACTAGAGGAGATCTCAGAGCAGATCCCT
rat_cdna human_5+3_corrected mus_cdna_5	TTTGGTTCAACAGAAAAATTGCTTGCAAGCTGAACAGAACTGCCACACCTAGCA CTTGGTTAATGCAAGACCAATTTCCTTGCAAGCTGAACAGAACTGCCACACATTAGTA
rat_cdna human_5+3_corrected mus_cdna_5	CATTACAGATCCAGTTTCCACTGATGCTCAATCGCTTACCAAGGGGGAGATGAGAG CATTACAGATCCAGTACTCCAGTGAATGCTCAAATCACTTACCAAGAGCAGAGATGAGGC
rat_cdna human_5+3_corrected mus_cdna_5	CGGAGAGACTCAAATGGACCATGATCCTGATGATGAACTGCCAACTGGAAAGCACTG CAGTGAACACAAATGGACTATGATTCAAGGGATAACATACTAAGCTGGAAACATACTG
rat_cdna human_5+3_corrected mus_cdna_5	TCCTGGTTGGCGGCACTATTGCCCTGAGCTGTCCAGGCCTAGGCCAGCCCTTCACCTCACT TCTGGTTAGGTGGAACCGTTGGCTGAACTGCCAGGCCTAGGAGACCCCCACCCACACG
rat_cdna human_5+3_corrected mus_cdna_5	TGGAATGGCTCTAGCTGATGGAGTAAAGTGAGAGGCCCTTACGTTAGCGAGGATGGGC TGGATTGGCTCTAGCTGATGGAGTAAAGTGAGAGGCCCTTATGTCAGTGAGGATGGAC
rat_cdna human_5+3_corrected mus_cdna_5	GAATCCTAATAGACAAAAATTGGAAGTTGGAACCTGCAGATGGCTGCGAGCTTGATGCGAG GGATCTAATAGACAAAGTGGAAATTGGAACCTCAGATGGCTGATAGTTTGACACAG
rat_cdna human_5+3_corrected mus_cdna_5	GCTTTACCTGCTAAGCAGCACTGATGCAGATGGGAATGTTCTACATACAGGATAA GGTATATCATGCTATAAGCAGCAATTATGATGATGCAAGTATTCTACCTATAGGATAA
rat_cdna human_5+3_corrected mus_cdna_5	CTGTGGTAGGGCTATGGAGAAGCAGCACATGACAGTGGAGTCAGCACACAGTGGTTA CTGTGGTAGAACCTTGTGCAAGCCATGAGGAATGGGATCAGCACACAGTTCTA
rat_cdna human_5+3_corrected mus_cdna_5	CGGGTGAGAGGCTCGACCTTOATGCTTTCCACGGGTGTTCCAGATGCTTCATTAAGCT TTGGTGAAACACTTGATCTCCATGOCATCTACTGGTATCCCAGATGCCCTCTATTAGCT
rat_cdna human_5+3_corrected mus_cdna_5	GGATTCTCCAGGGAAACACTGTTCTCTAGCCATCAAGAGACAGGCAAATTCTAAACA GGGTTATTCCAGGAAACATGTGCTCATCAGTCATCAAGAGACAGAAAGTTCTAAACA

rat_cDNA human 5+3 corrected mus_cDNA_5	ATGGGACCTTAAGAATATTACAGGTTAOGCCAAAGATCAAGGTCAATTACCAATGTGTGG ATGGCACATTAAAGAATATTACAGGTCACCCGAAAGACCAAGGTATTATCGCTGTGTGG
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCCAACCCATCAAGGGCCGACTTTTCCAGTTAAAGTTCAAGTCAAAAGAAAGGCC CAGCCAACCCATCAAGGGTTGATTTTGATTTCAAGTCAAGATGAAGGAC
rat_cDNA human 5+3 corrected mus_cDNA_5	AAAGGATGGTTGAGCATGACAGGGAGGCAGGTGGACTGGACTTGGAGAACCCAACTCCA AAAGGCCCTGGAGCATGATGGAGAACAGAGGGATCTGGACTTGTGAGTCACATCCCA
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGTTTCCCTTAAGCAGCCAGCACTTTGAAACTCTCTGCATCAAGCTTGTGACAGGGTCAG TTGCTCATCTTAAGGAGCCACCGAGGTGACAACTCGTACATCTGCTCTGATGAGTCACATCC
rat_cDNA human 5+3 corrected mus_cDNA_5	AGGCTGGAAAACAACTGTOCCGGTGTACATAGGAAGAACATAGAGACTTAATACATC AGGTTGAAAACACACCTCAAGCACAAGTAAAGAGGCACAACTATCGGAATTAAACACTCC
rat_cDNA human 5+3 corrected mus_cDNA_5	GGCGCGTGGGGATTCCACCGCTCGGGGCAITCAGGGAGCATAGGAGGCAGCTCOCTCTCT AGCGAGGTGGAGATTCAACACATGACCTTGTGGAGAATAGGAGGCATTCCCTCCCT
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCTCGGAGAATTGACCOGCAACGCTGGGCAACACTTCAGAAAAGCCAAAAGAATT CTGCTAGGAGAATTGACCCACAACATTTGGGGCACTGTGGAGAAAGCTAAAAAGATG
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTCGCCAAAAAGCAAGAAAATACCAACAGTAAAGCCAGTGCOACTGGCTGTTCCCTCG CTATGCCAGACAAGCGAGAAATACCAACAGTGTGGCCACCCOCAGTGGTCACCCAACTCC
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGAACTCACTGACCGAGAAAAGGATGCTCTGGCATGATCCAGATGAGAATTCA CAAAACATACTGGTGAAGAAGACGATTOCTCAGGCATGCTCGCTCTACATGAGGAATTAA
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGTTCTGAAACTAAGGCTCTGGTGTCCAGGAAGGTCAACACTGCTGACTCTGGAC TGGTCCCGGCCACTTAAGCTTGTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAA
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGTAATCATGGTTTATGACGAGTAAGCTCTGGCACAGAAGTCTCAGCTGTGATC CAATATCTGATAGTCATGACAAACATAAAATTATGGCACAGAAGCTCAGTGTGATT
rat_cDNA human 5+3 corrected mus_cDNA_5	CACAAACCTACATCTGAGCACCTTCCTGATTCAAAATTATTTAGTGTAAACAAAGCTA CACAAATACTACCACCTGAAGAACCCACAGATTTCAAACTGTCTACTGTGATTAACAACTA
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATAACACCAACC CAGCCATGTCAAAGATATAACCCACCATGTCAAGCAAAATCAAGGCAACCAACRATC
rat_cDNA human 5+3 corrected mus_cDNA_5	AAAACCCCATCATTATCTTCCATCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAACATC AACATTCATCCACTGTCTTCCACTGTCTACTGTGAGACTGTGAAATTTCAGGACTCTGACAA

rat_cdna	GGAGAGAGCATCTTCCAAAGTGCACACCCCTGAAACAGGGGAAACATGGCT
human_5+3_corrected	GAGGGAGAGGAAGAGGATTTCCAGTAACCCCAATAACAGTAAGGACTATGATC
mus_cdna_5	
rat_cdna	ACCTATGCCATACCAACACATAAGTAGCTTACCCAGCAAGCCAGTACAGTCTTGC
human_5+3_corrected	AAAGATGNTCAATGTCAAANATGCTTAGTAGCACCACCAACAAA-CTATTA-TTAG
mus_cdna_5	
rat_cdna	AGCCATAATCCAACAGAAAGTTATGGACCTCAGATAACCTTACAGGAGTCAGCAGAC
human_5+3_corrected	AGTCAGTAATACCCACAAATAGTCAT-CAGACATCTGTAAGAGAAGTGTGAAAC
mus_cdna_5	
rat_cdna	CTAGCAGTAGTGACATCTCTTCTCACACTACTGCAAGACCCCTAGCTTCTCCAGTCACCCCT
human_5+3_corrected	CCAGGCACATCTCACCTCTTCTCACACTACTCAAAATACTTAGCACCTCCACGTCCCTT
mus_cdna_5	
rat_cdna	CAGGTTCCACACCACACTGCCTGCTTATTTCACATTCAGAAACAACAATACAGGTA
human_5+3_corrected	CAGATCCACACACAGCTGCTCATCTCAGTTCCOGATCCCTAGANNAATAGTACAGTTA
mus_cdna_5	
rat_cdna	ACTTCCCTTGTCCAGGCACTGGGAAGAGAGGACAATTGGAGCAGAGGGAGAGTTA
human_5+3_corrected	ACATCCCCCTGTTAGACAGCTTCTGGGAGGCAGAGGAAATTGGGGAAGGGGGGGGTTA
mus_cdna_5	
rat_cdna	AAAACCCATAGAACCCCACTCTCGACGGCATAGACACAGGACTGTGAGGCCAGCAA
human_5+3_corrected	TCAGCCCATATAGAACTCGACTGCTCTGCGACGGCATAGATAACAGCATTTTCAGGTCAACRA
mus_cdna_5	
rat_cdna	TCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTCCAGGCCACAGAGTACCCCTGGGATGT
human_5+3_corrected	CCAGAGGTCTTCCTGAAAGCACTACTGCAATTCTCAGGCCACAGTGCTCATGTGACAT
mus_cdna_5	
rat_cdna	GCCACACATGTCTTCCGCAAGGGGCTCACAGTGGCTACTGCAAGCACTGTCAAGTCCRA
human_5+3_corrected	GTCTGTCCTGTCTTCCCAGGGAGGGCTCACACTGCCACAGCAGCATTTCTTTCRAA
mus_cdna_5	
rat_cdna	GTTCATCCCACAGTGCCCTCCCAAACATAATAATGTTGGGTCTAGCAGAAGACTCTA
human_5+3_corrected	GTGCGCTCCATCACCTTCCCAAAGCTGACATTGCTAGACTCCCATCAGAAGACTCTA
mus_cdna_5	
rat_cdna	CCACATGGTCAAGAACCACTGTTACTATTTAAGGACAAACAAATGAGATATTGAGA
human_5+3_corrected	CAACTCTAGTCAGAAATCCATTTACTACTCTGAGAACAAACCAAGTGAGANNNNGAAA
mus_cdna_5	
rat_cdna	TAATAACAAACCACTAACAAAATTCGGAGGGGAAAGTAACCACGTGATTCTCACGGAAAG
human_5+3_corrected	NNACAAACACCCACAATAAAATATTCAAGGACTGAAATTTCCTAAGTGAATCCAACTGGTG
mus_cdna_5	
rat_cdna	CAAGCATGACTTCTGCTCCAACATCTGTATCCCTGGGAAATCTCTGTAGACAATAGTG
human_5+3_corrected	CACTCATGACATATGCTCAACATCCATACCCATGGAAAAACTCACAAAGTAACAGCCA
mus_cdna_5	
rat_cdna	GTCACCTGAGCATGCTGGGACCATCCAAACTGGGAAAGATTCACTGGAAACACACAC
human_5+3_corrected	GTTACCCACCGTGTCTAGCCACCAATGAAGCTAAAGAGATTCACTGATTACATGTCAC
mus_cdna_5	

rat_cdna	TTCCCCAGCCCCCTCAGCAGACCCCCCAATACCAA-----	CAAGCAGCAGAAATTCTCAAGA
human_5+3_corrected	TTTCAGGGTGTATCACCAAGGCCACCATGACTATTATAGCCATTACAGGTTTCAGAA	
mus_cdna_5		
rat_cdna	GGAAAATCTCCCTGGCACAGAACCTTGTAAATAACCCAGAAGGGGGATGTTAAAGA	
human_5+3_corrected	GGAAAATCTCCCTGGCACAGAACCTTGTAAATAACCCAGAAGGGGGATGTTAAAGA	
mus_cdna_5		
rat_cdna	ATCCATATCAATTGGTTTACAAAAGAACCCAGCAGCAAGGTTCCCAAAATAGCTCTC	
human_5+3_corrected	ATCAACATAAAGTTAGTTACAAAAAGCAGCTGTGTTGCTTCTAAACATCTCTG	
mus_cdna_5		
rat_cdna	TTTTACCCACAG-GTCAGAGTTCCCCCTCAGATTCTACAACTCTCTTGACAAGTOCGCA	
human_5+3_corrected	CTTTACCCACAGAGACAGTTCCCCCTTCCATTTCACCCACACTTCAACAAGTGTGATG	
mus_cdna_5		
rat_cdna	CCAGCTCTGTCTACAAACATGGCTGCCACTCAGAACRAGGGCACTGAAGTAGTATCAGGT	
human_5+3_corrected	CAAAATTCTCATCTAATACCTTGACTACGGCTACCCACACTACGACCAAA--ACACACAA-T	
mus_cdna_5		
rat_cdna	GCCAGAAAGTCTCTCAGCAGGGAAAGAACGAGCCCTTCACCAACTCCTCTCCAG-TGCTTCC	
human_5+3_corrected	CCTGGAAAGTCTCCAAACAAAGAAGGAGCTTCCCTTC-CCACCCCTTAACCCATATGCTTCC	
mus_cdna_5		
rat_cdna	TAGCACCCTAACGAGAGATCTAATACATTAAACTTGTCAACGGAAACCCCCACAGT	
human_5+3_corrected	TAGTATTATAAGCAAGACTCAGTACAAAAGCATCATATCAACGCAAACAGCAGCAGC	
mus_cdna_5		
rat_cdna	GACAAGTCTACTGCTACTGCTCTGCTATTATGCTGAAACCAACGAACAAGATCCRA	
human_5+3_corrected	AACAACCTCTACCTTCCCTGCTCTGCTCTACTTATGAAACCAACAGAGAGACTCTAG	
mus_cdna_5		
rat_cdna	AGARGCAAAAGACCAAATTAAGGGGCTCG--GAAGAACGAAACAAAGCRAACACCCAC	
human_5+3_corrected	AGCACAAACATAAAAGAGAAGGACCTGAAAAGAACGGACTGACCCAAACATCTC	
mus_cdna_5		
rat_cdna	CCCCAGGCGGGTTCTGGCTATAGTGCATTACTCAGCTCTAACACACAGCTGATAACCCCTT	
human_5+3_corrected	TOCAGACCGAGTTCTGGCTTCTACCCACTGCTATGCGAACCTCTTNGCTTNNNN	
mus_cdna_5		
rat_cdna	GGCTTTCAGTCATTCCCAAGCACAGATGATGGTGGAAATGTAAGTCCAGTTGCTTATCA	
human_5+3_corrected	NGCATTCACTCATTCCCAAGCACAGAAACACAACTGGGATTTCAGCAATCAGTTCA	
mus_cdna_5		
rat_cdna	CTCAACAAACCTCTCTGGCCA--TAACTGAACTGTTGAG-AGTAC--ACCCAGAC	
human_5+3_corrected	TTCAAGAACTCTTAAATCTGACAGATGTGATTGAGAACTAGCCAAAGCAGTACTCAGAC	
mus_cdna_5		
rat_cdna	TTGGGAAATCACACAGCTTGGAAACAAAGCTTGTGAGCAAATCACAGGAGACTACCAC	
human_5+3_corrected	TTTGAAAGAGCACATGTCTCTGAAACACTTGTCCACAAATCACACCAAGGTACCCAC	
mus_cdna_5		
rat_cdna	ACTGAAAAGAGCCTCA--GACAC--ACCAACCAACTCCTCAGCAGTGGGGCGCCCG	
human_5+3_corrected	AACTAGGAAAGCAGCTCATAGACACTCACCCACCCATTCTGAGCAGCAGTACTCT	
mus_cdna_5		

rat_cDNA	AGTGGCCCACTCCCTCCCCACCCCTTTTACTAAGGGTGTGGTTACAGACAGCRAAGTCAC
human_5+3_corrected	AATGOCAGTTCCOCATCTCCCCCTCCCTTACTCAGAGGCACTGACAAACGTGGGAC
mus_cDNA_5	
rat_cDNA	ATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCATATGAATCTTCAGGACAA
human_5+3_corrected	TCCCATTTCCGGGCTTATGACAAATACAGTGGTCAGCTGCACGAATCTTCAGGACAA
mus_cDNA_5	
rat_cDNA	TACAGATCIGCAGCAACCCCTCAGCAGAGGGCTAGCCCCAATCCTGAGATCATAACTGGAAC
human_5+3_corrected	TCNNNNNNNCA---AATGCGAAGTTCA-----CNAATTGNGAACNNNACTGNNN
mus_cDNA_5	
rat_cDNA	CACTGACTCTCCCTCTAATCTGTTCCATCCACTCTGTGCCAGCACTAAGGGTAGATAA
human_5+3_corrected	NACTTCATCTACNTCTAATCTGTTACATCTACTCCCATGCCAGCACTAACACAGTTAA
mus_cDNA_5	
rat_cDNA	ACCRGAGAATTCTAAATGAAAGCCCTCTCCCTGGCCAGARORCAAATATCAGCTCAGTC
human_5+3_corrected	ATCACAGAAATTCCAATTAACCTCCATCTCCCTGGGCGAAATCCAATTGGCAGAAACC
mus_cDNA_5	
rat_cDNA	ATACTCGAAACCATTGAGAAGGGCAAAGGOCAGCTAAGCATGTCCCCCCTCCCTCAG
human_5+3_corrected	ATACTCGACATTGCTGAAAGGCGAAAGGCCAGAAGTAAGCATGTTGGCT-ACACAG
mus_cDNA_5	
rat_cDNA	-CCTTCCAGAGGCCAGCACTCATGCCCTCACACTGGAATACACAGAAGCATGCGAGAAAAGA
human_5+3_corrected	GOCTGTCGAGGCCACCACTCTGTTTCAAGATTGGGATGGACAGAAGAACACAAAGAGA
mus_cDNA_5	
rat_cDNA	GTGTTTTGATAAGAAACCTGGTCAA-AACC--CAACTTCAAACATCTGCCCTTACGTCT
human_5+3_corrected	GTGACTTTGATAAGAAACAGTTCAAGAAGCAACACTTCAAACCTCCCTCCCTTGTACT
mus_cDNA_5	
rat_cDNA	CTCTACCTAAGACTCTATTGAAAAAGCCAGAATAATTGGGGAAAGGCTGCAAGCTTTA
human_5+3_corrected	CTTTGTCTAGGTATATATTGAAAAGCCAGGATAGTGGGAGAAAAGCTGCAAGTTTA
mus_cDNA_5	
rat_cDNA	CAGTTCAGCTAATTCAAGACCTTTCTCCCTTGAGGGCTGTTGGAGACCCACTGCCCA
human_5+3_corrected	CTATTCCAGCTAACTCAAGATGCCCTTCTCCCTGTGAAGCTGTTGGAAATCCCTGCCCA
mus_cDNA_5	
rat_cDNA	TCATCCACTGGACCAGAGTTCTCAGGANTTGAATAATCCCAAGGGACACAGAAAAGCC
human_5+3_corrected	CCATTCAATTGGACCAGAGTNNNTCAAGGACTTGTATCTAAGAGGAACAGAATAGCA
mus_cDNA_5	
rat_cDNA	GGTTCACCGTCCTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAAGGACCGTG
human_5+3_corrected	GGGTTCAGGTTCTCCCAATGGTACCCCTGTCCATCCAGAGGGTGGAAATTCAAGGACCGTG
mus_cDNA_5	
rat_cDNA	GACAGTACCTG1GCTCTGCATTAACTCCACTGGCGTAGACCATTTCTATG1C1C1TTGT
human_5+3_corrected	GACAGTACTGTGTCCGCAATCTGTTGGCAGACAGGACCACTTCTGTCAACCTTGT
mus_cDNA_5	
rat_cDNA	CTGTGGTTTTACCGGCAAGGATTGGACAGACATGTCAGGAGATCACAGTTCACT
human_5+3_corrected	CTGTGGTTCTATCCTCCAGGATCTGGAGAGCGTACCAAAGAGATCACAGTTCACT
mus_cDNA_5	

rat_cDNA human_5+3_corrected mus_cDNA_5	TTGGAGTACTGTGGAACCTAAAGTGAGAGTGGAGGTATGCGAGGCTACGGTTTCTT CCGGAGCAGTGTGGAACTGAGTGAGCAGAGGTAAGCCCTACAGTACCT
rat_cDNA human_5+3_corrected mus_cDNA_5	GGATACTGCAACCAAACGGTGGTCTCAGAAACGCCAAGGGAGCAGAAAGGCTGGG GGATTCTGCACACCAAACAGTGTCTAGAATCTCOCAGGGAGTAGGCAGGCTGTGG
rat_cDNA human_5+3_corrected mus_cDNA_5	TAACACCTGATGGAACATGATCATCTATAATCTGAGTCTTTATGATGTTGGTTTACA TGACGGTGTGACGGAACATGGCTCCACAACTCTAGTATTTATGACGGTGGCTTTACA
rat_cDNA human_5+3_corrected mus_cDNA_5	AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCTGTTGGTTAAAGATAACAAGTCATCA AAATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCTGCTGGTTAAAGATAACAAGTCATTC
rat_cDNA human_5+3_corrected mus_cDNA_5	CAGCTCCCTGTCTTATAGAGCRAAGAGGCAAGCCATGTTGGGTTTTAGGTGGAA CAGCACCAACCTGTATCTAGAGCRAAGAGGCAAGTCATGTAAGGCACTTGGGTGAAA
rat_cDNA human_5+3_corrected mus_cDNA_5	GTTTGAAACTGCCCTGCACTGCAAAGAGGAACCTCCCGACGCTAGTGTTCAGTGGGTCTTT GTTTAAACTGCCCTGTACTGCAAAGGAACCTCCCTAGCCAGGCTTACTGGGTCTCT
rat_cDNA human_5+3_corrected mus_cDNA_5	ATGATGGGACTGAACCTAAACCATTCGAGTTGACTCACTTCAGATTTTCTTGATCCAA CTGATGGCACTGAAGTGAAACCTTACAGTTACCTTACAGTTGTTCTTATTTCAA
rat_cDNA human_5+3_corrected mus_cDNA_5	ATGGAACTCTGTATATAAGAACATGGCTCTTCAGTGAGGGGACTTATGAGTGCATTG ATGGGACTTTGTATATAAGAACCTGACCTCTTCAGACAGGGGACTTATGAAATGCATTG
rat_cDNA human_5+3_corrected mus_cDNA_5	CCACCAAGCTCTCAGGGCTCAGAGAGAAGGGTAGTGATTCTACTGTCGAAGAGGGAGAGA CTACCAAGCTTCCACTGGTTCGGAGOGAAGAGTAGTATGCTTACATGGAAAGAGOGAGTGA
rat_cDNA human_5+3_corrected mus_cDNA_5	CAATCCOCAGGATAGAAACTGOCCTCTCAGAAATGGACTGAGGTGAATTTGGTGAGAAAT CCAGCCCCAGGATAGAAAGCTGCACTCCAGAAAGGGACTGAGGTGAATTTGGGAGAAAT
rat_cDNA human_5+3_corrected mus_cDNA_5	TACTACTGAACCTGCTAGCTACTGGGATCCAAAGCCTAGAATAATCTGGAGGCTGOCAT TACTACTGAACCTGCTAGCCACTGGGAGGCCAAACCCCTAAATAATGTTGGAGGTACCAT
rat_cDNA human_5+3_corrected mus_cDNA_5	CCAGGGCTGTCACTGACCTGAGTGGCAAGAAATGGGAGGCCAAATCCAGCTACCCAAATG CCAAGGGCTGTGGTCACTGAGTGGCTACCTGAGTGGCTACCTGCTACCCAAATG
rat_cDNA human_5+3_corrected mus_cDNA_5	GATCCTGGTGGTGGCTAGTGACGGAAAAAGACGCTGGTACTACTTATGTGTGGCAA GATCCCTGTTATGATCAGTAACAGAAAAAGACAGTGGTGTCTACTTGTGTGGCAA
rat_cDNA human_5+3_corrected mus_cDNA_5	GAAACAAAATGGGAGATGACCTAGTCTGATGCATGTCGGCTGAGATTGACACCTGCGA GAAACAAAATGGGGATGATCTGATACTGATGCGATGTTAGCCTAAGACTGAAACCTGCCA

rat_cdna	AAATTGAAACAGAAGCACTATTAGAAGCAAGTGCCTCATGGAAAGATTTCCAACTGG
human_5+3_corrected	AAATTGAAACAGAAGCACTATTAGAAGCAAGTGCCTCATGGAAAGATTTCCAACTGG
<u>mus_cdna_5</u>	
rat_cdna	ACTGCAAGGCCTCTGGCTCCCTGIGGCTGAGGTATCTGGAGTTTGCTGATGGACAG
human_5+3_corrected	ATGCAAGCTCCGGCTCCCTGAGGTGAGGATATCTGGAGTTTGCTGATGGACAG
<u>mus_cdna_5</u>	
rat_cdna	TGCTCAACAAATGTAGCCAAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCC
human_5+3_corrected	TGATCAACAAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATAACCC
<u>mus_cdna_5</u>	
rat_cdna	TCCACAATGGAACCTTGATTTCAACAACGTTGGATGGCAGAGGAAGGAGTTATATCT
human_5+3_corrected	TCAACAATGGAACCTTATACTTCACAACAAAGTGGGTAGGGAGAGGAGTTATACTT
<u>mus_cdna_5</u>	
rat_cdna	GCTCTGCCAGAACACCTTGGAAAGATGAGATGAAAGTCCACCTAACAGTTCAACAG
human_5+3_corrected	GCTATGCCAGAACACCTTGGAAAGATGAAATGAAAGTCCACCTAACAGTTCAACAG
<u>mus_cdna_5</u>	
rat_cdna	CCATCCCACGGATAAGGAAACCTAACAGGCTAACAGACCAACATGAGGCTCAGGCTGGAGAACAG
human_5+3_corrected	CTGCTCCCCGATAAGGCAAGATGAAACAAACAAACAAAGAATCAAGCTGGAGAACAG
<u>mus_cdna_5</u>	
rat_cdna	CTGTCCTTGACTGCGAGGTCACTGGGAACCGAACCGAACATGTATTTGGTTGCTGCTT
human_5+3_corrected	CTGTCCTTGACTGCGAGGTCACTGGGAACCGAACATGTATTTGGTTGCTGCTT
<u>mus_cdna_5</u>	
rat_cdna	CCAACAATGTCATTTCAATTCTCAATGACAGGTTACATTTCAAGCCAAATGAAAGTCTTGT
human_5+3_corrected	CCATGACATGATTTCTTCATTGATAGGTACACATTTCAAGCCAAATGGTCTTTGA
<u>mus_cdna_5</u>	
rat_cdna	CCATCCATAAAGTGAACACACTTGACTCTGGGACTATGTGTGCGTAGCTCAGAAATCTA
human_5+3_corrected	CCATCAACAAAGTGAACACACTGCTGATCTGGAGAGTACSTATGTGTAGCCAAATCTA
<u>mus_cdna_5</u>	
rat_cdna	GTGGGGATGACACTAAAGACATACAAACTGGACATTGCTCTAAACCTCCATTATCAATG
human_5+3_corrected	GTGGGGATGACACCAAAATGTACAACACTGGATGTGGCTCTAAACCTCCATTATCAATG
<u>mus_cdna_5</u>	
rat_cdna	GCCTGTATGCAACAAAGACTGTTATTAAAGCCACAGGCCATTGGCACTCCAAAAAAACT
human_5+3_corrected	GTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTTCAAAACACT
<u>mus_cdna_5</u>	
rat_cdna	TTGACTGCAGAGGAGATGGGATCCCATCTTCCGAGGTCAAGTGGATTATGCCAGGAATA
human_5+3_corrected	TTGACTGCAGAGCTGAAGGACACCCATCTGAAACTGTGAGTGTGGATCAAGCCAGCAATA
<u>mus_cdna_5</u>	
rat_cdna	TTTCCCTCCAGCTCCATCTTGGAGAGTCACGGTOCATCCAAATGGAACTTGG
human_5+3_corrected	TTTCCCTCCAGCCATCTATGAAAGCAGAGATCACACTCCATAAAATGGAACTTGG
<u>mus_cdna_5</u>	
rat_cdna	AGATGAGGAACCTCCGGCTTCTGACTCTGGGACTTCAGATTGAGGACTTATCTGTGTCGGAGGAGG
human_5+3_corrected	AAATTAGGAATGTGAGGCTTCAGATTGAGGACTTATCTGTGTCGGCCGAAATGAAG
<u>mus_cdna_5</u>	

rat_cDNA	GAGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAGAAGACCAACAT
human_5+3_corrected	GTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGACCGACAT
mus_cDNA_5	
rat_cDNA	TCAGAAACCCATTCAACGAAAAAGTCATGCCAAGCTGGCAAGCCGTAAGCAGTGAACCT
human_5+3_corrected	TTAGAAATCCATTAAATGAAAAATAGTTGCCAGCTGGAAAGTCCACAGCAATGAAATT
mus_cDNA_5	
rat_cDNA	GCTCTGTGGATGGGAAACCCCCCACCTGAAATTACCTGGATCTTACCTGAOGGCACACAGT
human_5+3_corrected	GCTCTGTGGATGGTAACCCACCCACCTGAAATAATCTGGATTTACCAAATGGCACACGAT
mus_cDNA_5	
rat_cDNA	TTGCTAACAGACCCACAATTCCCCGTACTGATGGCAAGGCAATGGCTCTCATCTTT
human_5+3_corrected	TTTCATTGGACCAAAAGTTACAGTATCTGATAGCAAGCAATGGTTCTTTATCATTT
mus_cDNA_5	
rat_cDNA	ACAAAGCAACTCGAACAAAGTCAGGAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT
human_5+3_corrected	CTAAACAACTCGGAGATGCAGGAAATAATCGCTGTGCAGCTAGGAATAAGGTTGGCT
mus_cDNA_5	
rat_cDNA	ACATCGAGAAACTCATCCGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATAAGAAC
human_5+3_corrected	ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAGCCAGTTATCTTACCTATGCCAC
mus_cDNA_5	
rat_cDNA	CAGGGATGGTGAAGAGCGTCAGTGGGGAAACGTTATCACTGCATTGGCTCTGATGGGA
human_5+3_corrected	CAGGGACAGTAAAAGGCATCACTGGAGAAATCTCTATCACTGCATTGGCTCTGATGGGA
mus_cDNA_5	
rat_cDNA	TCCCCAAGCCAAATGTCAGTGGACTCACCGGGTGGCCATGTAATGACAGGCTCAAG
human_5+3_corrected	TOCTTAAGCCAAATATCAATGCGCTATGCCAAGTGGTTATGAGTAAACAGGCTCAAA
mus_cDNA_5	
rat_cDNA	TGGATGGAAAAATACATACTGCATGAAAAATGCCAOGCTGGTCATCAAAGCAACACAGCTC
human_5+3_corrected	TTATGGGAAATACATATTGCAATGACAATGCCACCTAGTCATTAAGAAGCAACAGCTT
mus_cDNA_5	
rat_cDNA	ACGAACGAGGAATTATATCTGTAGGGCTAAACAGTGTGGCAGGGCACTTATTAGCG
human_5+3_corrected	ATGACAGAGGAAACTATATCTGTAAAGGCTAAATAGTGTGGTCATACACTGATTACIG
mus_cDNA_5	
rat_cDNA	TGTCAGTGTGGTGTGGCTAACCTCCCGAATGATAACTAACCTACCCAGGACAGTC
human_5+3_corrected	TTCCAGTAATGATTGTAGCTAACCTCCCGAATTAACAAATGTCACCCAGGAGTATTG
mus_cDNA_5	
rat_cDNA	TCAGGAGGACAGGGCAACCGCTGGCAGCTCCACTGTGTGGCTTGGGAATCCCAAGCCAA
human_5+3_corrected	TCACCAGGACAGGGCAGCCCTTCAGCTTCACTGTGTGGCTTGGGAATCCCAAGCCAG
mus_cDNA_5	
rat_cDNA	AAGTCACCTGGAGACGCCAAGACACTCCGTCTCAAAGCAACAGCAGAAAACCCC
human_5+3_corrected	AAATCACTGGAGATGCTGTGACCACTCCCTCTCAACGCCAAGTAAGAGAGGACAC
mus_cDNA_5	
rat_cDNA	ATAGAAGTGAAGTGCTTCAACCAACAGGTACGCTGGTCATTGAGAATCTCCAAACCTCGG
human_5+3_corrected	ATGGAAGTGAAGCAGCTTCACTTACAGGTACCCCTAGTCATTGAGAATCCAAACCTCGG
mus_cDNA_5	

rat_cDNA	ATTCGGAGTCTATAAGTGCAGAGCTCAGAACCTACCTGGGACTGATTACGCCAACACTT
human_5+3_corrected	ATTCGGGATATAACAAATGCCRGCAAAGAACCCACTTGGTAGTGATTATGCCAGCAACGT
mus_cDNA_5	
rat_cDNA	ACATCCAGGTACTCTGACAGGAGGGGGAGACTAAATTCAACAGAAGTCACATOCACA
human_5+3_corrected	ATATTCAGTAATCTGACATGAA-----ATATAAAGTCAACAA-----CATCTGGCA
mus_cDNA_5	
rat_cDNA	GGGTTTATTTTGGAAAGAAGTTAACAAAGGCAGCCATAGGCATGTAAATGAGTCCTGA
human_5+3_corrected	GAATTTATTTTGGAAAGAAGTTAACAAAGGCAGCCATAGGCATGTAAATGAAATTGAA
mus_cDNA_5	
rat_cDNA	ATACATTACAGTATTAATTTACATGGACATGCCA---TGA---GACTTGAAATGAA
human_5+3_corrected	ATACATTACAGTATTAATTTACATGGACATGCCAAATAAAGGACTTGAAATGAA
mus_cDNA_5	
rat_cDNA	GCATTGTGAACGTGAA---ACCGAGCTCTG---TGGCTCTCAAACCPAAACTCTAACCTAA
human_5+3_corrected	GCATTATGAACGTGATGACTGTGATTTATTTAACATGCCAAACAAACTTTAACCTAA
mus_cDNA_5	
rat_cDNA	GGCCTTTGATTTGCCAACAAATAACAAACATTAAAGGAGAAAAATGATCCACTAC
human_5+3_corrected	GGCCTTTATTTGCCAACAAATAACAAACATTAAAGGAGAAAAATGATCCACTAC
mus_cDNA_5	
rat_cDNA	GAATAACAAACGGCTAATGCCACCTGAATTCT---CAGTAAAAAGACCTTTCTCTGGCTAAC
human_5+3_corrected	AAAATAACAAATGGCTATGTACCTGAATTTCAGTAAAAAA---TGAACCTTCTAAT
mus_cDNA_5	
rat_cDNA	AGTTGCCAGCTGCCCTCGTGTCTGTTCCACCAATGTCAACAAACATGCCACACAGGGTGA
human_5+3_corrected	A---CCAGTTGCCCTAGTGTCCACCTCTATCAATGTTAACAGCATGGCACTCAG---A
mus_cDNA_5	
rat_cDNA	ATGGAGTCACGGGAAAGATTAAAGTTTGCCTCTGTTAACATCCTAACATGTCACAAATTC
human_5+3_corrected	ACAGAGACATGGAAATATTAAATCTGC-----ATCT---ATGTATAAAATTTT
mus_cDNA_5	
rat_cDNA	TGTCTGGTTATAAAACATTTCGATAAAACCGAAAAAATTTTTTTTTTGTCTGGCTAAACCTACAGAAATAAG
human_5+3_corrected	TGT---GGTTTATAAAATTTTTGCTAAACCTACAGAAATAAG
mus_cDNA_5	
rat_cDNA	AAAA
human_5+3_corrected	---
mus_cDNA_5	

(rat_cDNA: SEQ ID NO:7)

(human_5+3 corrected: SEQ ID NO:8)

(mus_cDNA_5: SEQ ID NO:9)

Figure 15

rat	MQVRGREVSGLLISLTAVCLVVTGSRACPRRCACIVPTEVHCTFRYLTSPIDGIPANVE
human_5+3_corrected	MKVKGRRGITCLLVSPAVICLVTATPGKACPRRCACIYMPTEVHCTFRYLTSPIDGIPANVE
mouse_5_corrected	MOKRGREVSCLLISLTAVCLVVTGSRVCPRRCACIVPTEVHCTFRDLSIPD-GPANVE
rat	RINLGYNSLTRLTENDFDGLSKLELLMLHSGNGIHRVSDKTFSGLQSLQVLMKSYNKVQII
human_5+3_corrected	RINLGYNSLVRLMETDFSGLTKELELLMLHSGNGIHTIPDKTFSDIQLQVLMKSYNKVRL
mouse_5_corrected	RVNLGYNSLTRLTENDFDGSLRLELLMLHSGNGIHRVSDKTFSGLQSLQVLMKSYNKVQII
rat	RKDTFYGLGSILVRLLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLLTKLHPDTFVSLSYLQ
human_5+3_corrected	QKDTFYGLRSLLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNRLLTKLHPDTFVSLSYLQ
mouse_5_corrected	EKDTLYGLRSLLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLLTKLHPDTFVSLSYLQ
rat	IFKTSPIKYLFLSDNFTLSPKEMVSYMPNLESLYLHGNPWTCDCHLKWLSWMOGNPDI
human_5+3_corrected	IFKISPIKFLYLSDNFTLSPQEMVSYMPDLSLYLHGNPWTCDCHLKWLSWMOGNPDI
mouse_5_corrected	IFKTSPIKLYLYDNF-TSLPKEMVSSMPNLESLYLHGNPWTCDCHLKWLSWMOGNPDI
rat	IKCKKDRSSSSPQOCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLSVTOEDNG
human_5+3_corrected	IKCKKDRSPSSAQCPLCMNPRTSKGKPLAMVAAAQCAKPTIDSSLSKSLSLIEDSS
mouse_5_corrected	-----
rat	SASTSPQDFIEPFGSLSLNMTXSGNKADMVCSIQKPSRTSPTAFTEENDYIMINASPT
human_5+3_corrected	SAFISPQGEMAPFGSLLTINMTDQSNEAMVCSIQKPSRTSPIAFTEENDYIVLNTSFST
mouse_5_corrected	-----
rat	NLVCSDYDYNHIQPVWQILALYSDSPLILERKPKQLTETPSLSSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIDYGHIQPVWQILALYSDSPLILERSHLLSETPQLYYKQVAPKPEDIFTNIEA
mouse_5_corrected	-----
rat	DVRADPFWFQQEKRIVLQLNRTATTILSTLQIQTSTDAQIALPRAEMRAERLKWTMILMMNN
human_5+3_corrected	DLRADPSWIMQDQISLQLNRTATTSTLQIQTQSSDAQITLPRAEVMPVHKWTMISRDNN
mouse_5_corrected	-----
rat	PKLERTVLVGGTIALSCPGKGDPSPHLENLLADGSKVRAPIVSEDRGILIDRKNGKLELOM
human_5+3_corrected	TKLEHTVLVGGTVGLNCPGQGDPPTPHDVWLLADGSKVRAPIVSEDRGILIDRKNGKLELOM
mouse_5_corrected	-----
rat	ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDGVQHVVGETLDLPCLSSTGV
human_5+3_corrected	ADSFDGTGVYHCISIISYDDADILTYRITVVEPLVAYQENGIIHHTVFIGETLDLPCHSSTGI
mouse_5_corrected	-----
rat	PDASISWILPGNTVFSQPSRDRQIILNNGTIRLQLVTPKDOGHYQCVAAAPSGADFSSFKV
human_5+3_corrected	PDASISWIPGNNVLYQSSRDKVVLNNGTIRLQLVTPKDOGHYQCVAAAPSGGVDFLIFQV
mouse_5_corrected	-----
rat	SVQKKQRMVEHDREAGGSGLGEPNSSVSLKQPAISLKLSSAALTGSEAGKRVSGVHRKNK
human_5+3_corrected	SVRMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEGKHTSSTSKRHN
mouse_5_corrected	-----
rat	HRDLIHRGGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKRNQVPKQENTTVKPV
human_5+3_corrected	YRELTQRRGGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKRNQVPKQENTTVKPV
mouse_5_corrected	-----

rat	PLAVPLVELTDEEKDASGMIPPDEEFMVLATKASGVPGRSPTADSGPVNHEGPMTSIASGT
human_5+3_corrected	PVVTOLPNIPGEEDDSQMLALHESEFMVPAKALNLPARTVTADSRTIISDPMTNINYGT
mouse_5_corrected	
rat	EVS-TVNPOQTLQSEHLPDFKLFSTVNGTAVTKSMNPSIASKIEDTTNQNPIIIIFP--SV
human_5+3_corrected	EFSPVVNSQILPPEEPTDFKLSTAIKTTAMSKNINPTMSSQIQGTTNQHSSTVFPILLGA
mouse_5_corrected	
rat	AEIRDASAQAGRAS--SQSAHPVTGGRNATYGHNTNTYSSFTSKASTVLOPINPTESYGPQI
human_5+3_corrected	TEFQDSIDQGRGRHIFQSRPPITFVRLMIKVNVKMLSTTNKL--LLESVNTTNSH--QT
mouse_5_corrected	
rat	PITGVSRPSSSDISSETTADPFSSSHPSGSHTTASSLFHIPRNNTGNPPLSRHLGRERT
human_5+3_corrected	SVREVSEPRHNHFYSHITQTLSTSTFPSPHTAHSQFPIPRNS-TVNIPFLRFGRQRK
mouse_5_corrected	
rat	IWSRGRVENPHRTPVIRRERHRTVRPAIKGPANKVSVQVPATEYPMCHTCPASAGLTV
human_5+3_corrected	IGGRGRIISPYRTPVIRRHYSIFRSTTRGSSEKSTTAFSATVNLVTCLSCLPRERLTTA
mouse_5_corrected	
rat	TAALSVPSSSHSSALPKNNVGVIAEESTTVVKKPLLFKDNQVNDIEIITTTTKYSGGES
human_5+3_corrected	TAALSFPSAAPITFFKADILARVSEESTLVQNPILLLENKF--SVEKTTPTIKYPRTEI
mouse_5_corrected	
rat	NHVIPTEASMTSAPTSVSLGKSPVDSGHLSPGTTQGKDSVETPLPSPLSTP--SIP
human_5+3_corrected	SQVTPGAVMTYAPTSIPMEKTHKVNASYPRVSSSTNEAKRDSVITSSLGAITKPPMTII
mouse_5_corrected	
rat	TSTKFSKRKTPLEQIFVNNOKEGMLKNPYQFGLOKNPAAKLPKIAPLLPTGQSSPSDST
human_5+3_corrected	ATRFSRRKIPWQONFVNHHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFT
mouse_5_corrected	
rat	TLLTSPPEALSTMAATONKGTEVVSGARSLSAKKQ-PFFNSSPVLPSLICKRSNTLNF
human_5+3_corrected	TISTSVQIPSNLTTAHHTTTKTHNPG-SLPT-KKELPFPPLNPMLPSIISKDSSTKSI
mouse_5_corrected	
rat	LSTETPT-VTSPTATASVIMSETQRTSKEAKDQIG-P-RKQRRNNANTTPRQVSGYSAV
human_5+3_corrected	ISTQTAIPATTPTFFASVITYETQTERSRAQTIQREQEPOKENRTDPNISPDQSSGFTTP
mouse_5_corrected	
rat	SALTTADTPLAFSHSPQDDGGNVSAVAYHSTTS--LIAITELFEKITYQTLGNTTALETT
human_5+3_corrected	TAMTP--PALAFTHSPPENTGIGSSTISFHSRTLNLTDVIEELAQASTOTLKSTIASETT
mouse_5_corrected	
rat	LLSKSQESTTVKRAS-DTP-PFLLSSGAPPVPTPSPPPFKGVVTDISKVTSAFQMTSNRV
human_5+3_corrected	LLSKSHQSTTRKASLDTPIPPFLSSSATIMPVPISPPFQRAVTDTREGDSHEFLMTNTV
mouse_5_corrected	
rat	VTIYESSRHENTDLQOPSAEASPNEITGTTDSPSNLFPSSTSVEALRVDKPQNSKWKPS
human_5+3_corrected	VKLHESRHN--LQMPSSOLEP-----LTSSTSNNLMSTPMPALTTVKSQNSKLTPSP
mouse_5_corrected	
rat	WPEHKYQLKSYSETIERGKRPAVSMSPHLSLPSASTHASHWNTQKHAEKSVDKKPGQNP
human_5+3_corrected	WAEYQFWHKPYSDIAEKGKKEPEVSMILATTGLSEATLTVSDWDGOKNTKKSDFDKKPVOEA
mouse_5_corrected	

rat	-TSHLPYVSLPKTLLKKPRIIGGKAASFTVPA	NSDVFPLCEAVG	OPLP	IIH	WTRVSSGX																																								
human_5+3_corrected	TSKLLPFD	LSR	YIFPEK	PRTIVGGKAASFTI	PANSDAFLP																																								
mouse_5_corrected	TSKLLPFD	LSR	YIFPEK	PRTIVGGKAASFTI	PANSDAFLP																																								
rat	EISQGTQ	SRFHVL	PNGT	LSIQRV	SIQDRGQYLC	CSAFNPLGV	DHF	EVLSVV	YPARILD																																				
human_5+3_corrected	LSRGNQ	NSR	VQVLP	PNGT	LSIQRV	EI	QDRGQYLC	CSA	NLFGT																																				
mouse_5_corrected	LSRGNQ	NSR	VQVLP	PNGT	LSIQRV	EI	QDRGQYLC	CSA	NLFGT																																				
rat	RHVKEITV	HFG	STVELK	CRVE	QMP	PTV	WILAN	OTVV	SETAKG	SRKV	WV	TPDGT	LI	YN																															
human_5+3_corrected	RTKEITV	HSG	STVELK	CRA	GRPS	PTV	WILAN	OTVV	SESSO	GROAVV	ITV	DGT	LI	VLEN																															
mouse_5_corrected	RTKEITV	HSG	STVELK	CRA	GRPS	PTV	WILAN	OTVV	SESSO	GROAVV	ITV	DGT	LI	VLEN																															
rat	LSLYD	RGFT	YKCV	ASNP	SGQD	SLLV	KI	QV	ITAPPV	II	EQK	RAIVG	V	LGSL	KLPCT	AKGT																													
human_5+3_corrected	LSIYD	RGFT	YKCV	ASNP	GGQD	SLLV	KI	QV	IAAPPV	II	EQR	QVIVG	V	WGES	KLPCT	AKGT																													
mouse_5_corrected	LSIYD	RGFT	YKCV	ASNP	GGQD	SLLV	KI	QV	IAAPPV	II	EQR	QVIVG	V	WGES	KLPCT	AKGT																													
rat	POPSV	EWV	LYDGT	ELK	PLQ	LTH	SRF	LYP	MNGT	LYI	TRSI	APS	VRG	T	ECIAT	SSSG	SERRV																												
human_5+3_corrected	POPSV	EWV	LYDGT	EVK	PLQ	TNS	KL	FL	FSN	GT	LYI	RNL	ASS	DRG	T	ECIAT	SS	SERRV																											
mouse_5_corrected	POPSV	EWV	LYDGT	EVK	PLQ	TNS	KL	FL	FSN	GT	LYI	RNL	ASS	DRG	T	ECIAT	SS	SERRV																											
rat	VIL	TVEE	GETI	PRIET	ASQ	QNT	TEV	NL	GEK	LL	LNCS	ATGDP	KPRI	IWR	RLPS	KAVID	QWHRM																												
human_5+3_corrected	VML	TMEER	RTSP	RIEA	ASQ	QRT	TEV	NF	GD	LL	LNCS	ATGEP	KPQ	IWR	RLPS	KAVID	QWHRM																												
mouse_5_corrected	VIL	TVEE	GETI	PRIET	ASQ	QNT	TEV	NL	GEK	LL	LNCS	ATGEP	KPQ	IWR	RLPS	KAVID	QWHRM																												
rat	GSR	IHV	PNG	SLV	VGS	VTE	KDAG	DYL	CVAR	NKMG	DDLV	LMH	VRL	IRI	TPAKI	EOKO	YFKKQ																												
human_5+3_corrected	GSI	IHV	PNG	SLV	FIG	S	TEK	DG	VYLC	VAR	NKMG	DDLV	LMH	VRL	IRI	TPAKI	EOKO	YFKKQ																											
mouse_5_corrected	GSR	IHV	PNG	SLV	FIG	S	TEK	DG	VYLC	VAR	NKMG	DDLV	LMH	VRL	IRI	TPAKI	EOKO	YFKKQ																											
rat	VLEGKDF	QV	DCKA	SGSP	VPEV	WS	SLP	DGT	V	NNV	QADD	SGY	RTK	RT	YF	NNV																													
human_5+3_corrected	VLEGKDF	QV	DCKA	SGSP	VPEI	WS	SLP	DGT	M	NNN	QADD	SG	RT	RT	YF	NNV																													
mouse_5_corrected	VLEGKDF	QV	DCKA	SGSP	VPEI	WS	SLP	DGT	M	NNN	QADD	SG	RT	RT	YF	NNV																													
rat	GM	REE	GDY	ICSA	QNT	LG	DEM	KV	H	LT	V	TA	PRI	OSY	CT	MLR	AGET	AVLD	CEV	TGEP																									
human_5+3_corrected	GM	REE	GDY	ICSA	QNT	LG	DEM	KV	H	LT	V	TA	PRI	OSN	KT	MLR	AGDT	AVLD	CEV	TGDP																									
mouse_5_corrected	GM	REE	GDY	ICSA	QNT	LG	DEM	KV	H	LT	V	TA	PRI	OSN	KT	MLR	AGDT	AVLD	CEV	TGDP																									
rat	KPN	V	FW	LLP	SNN	VI	FS	ND	RT	PHAN	RT	LSI	HKV	KPL	DSG	DY	VCVA	ONP	SG	DD	T	TY	KLD																						
human_5+3_corrected	KPK	I	FW	LLP	SNN	VI	FS	ND	RT	PHAN	RT	LSI	HKV	KPL	DSG	DY	VCVA	ONP	SG	DD	T	TY	KLD																						
mouse_5_corrected	KPN	V	FW	LLP	SNN	VI	FS	ND	RT	PHAN	RT	LSI	HKV	KPL	DSG	DY	VCVA	ONP	SG	DD	T	TY	KLD																						
rat	IVSKP	PL	LING	Y	AKT	V	IK	A	TR	HS	KKY	FD	CRA	DG	IP	SS	QV	W	IMP	G	P	FG	SR																						
human_5+3_corrected	IVSKP	PL	LING	Y	AKT	V	IK	A	TR	HS	KKY	FD	CRA	DG	IP	SS	QV	W	IMP	G	P	FG	SR																						
mouse_5_corrected	IVSKP	PL	LING	Y	AKT	V	IK	A	TR	HS	KKY	FD	CRA	DG	IP	SS	QV	W	IMP	G	P	FG	SR																						
rat	VT	VHP	NGT	LE	MRN	I	RL	S	AD	F	C	V	R	S	EG	GE	S	V	V	W	Q	L	E	M	R	P	N	E	K	V															
human_5+3_corrected	VT	VHP	NGT	LE	MRN	I	RL	S	AD	F	C	V	R	S	EG	GE	S	V	V	W	Q	L	E	M	R	P	N	E	K	V															
mouse_5_corrected	VT	VHP	NGT	LE	MRN	I	RL	S	AD	F	C	V	R	S	EG	GE	S	V	V	W	Q	L	E	M	R	P	N	E	K	V															
rat	QAG	KP	VAL	NC	SD	GN	PP	PE	I	W	PDG	T	Q	F	AN	R	P	Y	I	M	G	N	S	L	I	Y	K	A	T	R	N	K	S	G											
human_5+3_corrected	QLG	KST	ALN	CS	VD	GN	PP	PE	I	W	PDG	T	Q	F	AN	R	P	Y	I	M	G	N	S	L	I	Y	K	A	T	R	N	K	S	G											
mouse_5_corrected	QAG	KP	VAL	NC	SD	GN	PP	PE	I	W	PDG	T	Q	F	AN	R	P	Y	I	M	G	N	S	L	I	Y	K	A	T	R	N	K	S	G											
rat	RCA	ARN	KV	GY	IE	K	L	I	L	E	I	G	K	P	V	I	L	T	Y	E	P	G	M	V	S	G	E	P	L	S	I	H	C	V	S	D	G	I	P	K	P	N	I	W	M
human_5+3_corrected	RCA	ARN	KV	GY	IE	K	L	I	L	E	I	G	K	P	V	I	L	T	Y	E	P	G	M	V	S	G	E	P	L	S	I	H	C	V	S	D	G	I	P	K	P	N	I	W	M
mouse_5_corrected	RCA	ARN	KV	GY	IE	K	L	I	L	E	I	G	K	P	V	I	L	T	Y	E	P	G	M	V	S	G	E	P	L	S	I	H	C	V	S	D	G	I	P	K	P	N	I	W	M

rat	GGHVIDRPQVDGKYILHENGTLVIKATTAHDQNYICRAQNSVGQAVIEVSVMVVAYPPR
human_5+3_corrected	SGYVVDRPQINGKYILHEUNGTLVIKEATAYDRGNYICKAQNNSVGETLITVPVMTIVAYPPR
mouse_5_corrected	

rat	IINYLPRNMLRRTGEANOLHCVALGIPKPKVITWETPRESILSKATARKPHRSEMLHPQGT
human_5+3_corrected	ITNRPPRSIVTRTGAAPOLHCVALGVPKPEITWEMPDHSILLSTASKERTHGSEQLHLQGT
mouse_5_corrected	

rat	LVIONLQTSDSGVYKCRAQNLLGTDYATTYIQV
human_5+3_corrected	LVIONPQTSDSGIYKCTAKNPLGSDYATTYIQV
mouse_5_corrected	

(rat: SEQ ID NO:10)

(human_5+3_corrected: SEQ ID NO:11)

(mouse_5_corrected: SEQ ID NO:12)

Figure 16

rat	SRIHVYPNGSLVVGSVTEKDAGDYLCVARNKMGDDLVIMHVRRLT PAKIEQKQYFKKQV
human_5+3_corrected	SWIHVYPNGSLFIGSVTEKDAGDYLCVARNKMGDDLVIMHVRRLT PAKIEQKQYFKKQV
rat	*****;*****;*****;*****;*****;*****;*****;*****;*****;
human_5+3_corrected	LHGKDFQVDCDKASGSPVPEVSWSLPDGTIVLNNAQADDSGYRTKRYTLEFHNGTLYFNNG
rat	LHGKDFQVDCDKASGSPVPEIWSLSPDGTMINNAQADDSGYRTKRYTLEFHNGTLYFNNG
human_5+3_corrected	*****;*****;*****;*****;*****;*****;*****;*****;*****;
rat	MAEEGDYICSAQNTLGKDEMVKHLTVLTAIPRIRQSYKTTMRLRAGETAVLDCEVTGEPK
human_5+3_corrected	VAEEGDYTCYAQNTLGKDEMVKHLTVITAAPIRIRQSYKTTMRLRAGETAVLDCEVTGDPK
rat	***** * *****;*****;***** *;***** *;*****;*****;*****;*****;
human_5+3_corrected	PNVFLLPSNNVISFSNDRFTIHNRTLSIHKVKPLDSGYVCVAQNPSGDDTKYKLDI
rat	PKIFWLLP9NDM16FSIDRYTFHANGSLTKLLDSGEYVCVARNPNSGDDTKMYKLDV
human_5+3_corrected	*****;*****;*****;*****;*****;*****;*****;*****;*****;
rat	VSKPPLLINGLYANKTVIKATAIRHSKKYFDRADGIPSSQVTWIMPGNIFLPAPYFGSRV
human_5+3_corrected	VSKPPLLINGLYNRTVIKATAVRSKHKFDRAEGTPSPFVWIMPDNIFLTAPYYGSRI
rat	*****;*****;*****;*****;*****;*****;*****;*****;*****;
human_5+3_corrected	TVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLLEMRLRRPTFRNPFNEKVIAQ
rat	TVHKNGTLEIRNVALSDSADFTCVARNEGGESVLVVQLEVLLEMRLRRPTFRNPFNEKVIAQ
human_5+3_corrected	*****;*****;*****;*****;*****;*****;*****;*****;*****;*****;
rat	AGKPVALNCSDVGNPPPEITWILPDGTQFANRPHNSPYLIMAGNGSLILYKATRNKSGKYL
human_5+3_corrected	LGKSTALNCSDVGNPPPEIWIWLPNGTRFSNGQSYQYLIASNGSFTIISKTREDAGKYL
rat	*****;*****;*****;*****;*****;*****;*****;*****;*****;
human_5+3_corrected	CAARNKVGYIEKLILLEIGQKPVILTYEPGMVKSVSGEPLSLHCVSDGIPKPNVKWTPG
rat	CAARNKVGYIEKLVILLEIGQKPVILTYAPGTVKGISGESLSLHCVSDGIPKPNIKWTPS
human_5+3_corrected	*****;*****;*****;*****;*****;*****;*****;*****;*****;
rat	GVITDRPOVDGKYILHENGTIVIKATAHDQGNYICRAQNSVGQAVISVSVMVAYPPRI
human_5+3_corrected	GYVVDRPOINGKYILHDNGTIVIKEATAVDRGNYICKAQNNSVGETLITVPPMIVAYPPRI
rat	*****;*****;*****;*****;*****;*****;*****;*****;*****;
human_5+3_corrected	INYLPRNMLRRIGEAMOLACVALGIPPKVWETPRHSILSKATARKPRSEMLHPQGTL
rat	TNRPSSIVTRGAAFOLHCVALGVPKPEITWEMPDHSSLSTASKERTHGSEQLHLOGTL
human_5+3_corrected	*****;*****;*****;*****;*****;*****;*****;*****;*****;
rat	VIONLQTSDSGVYKCRAGNLLGTDYATTYIQVL
human_5+3_corrected	VIONNPQTSDSGVYKCTAKNPLGSDYAATYIQVI
rat	*****;*****;*****;*****;*****;

(rat: SEQ ID NO:13)

(human_5+3_corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTACILVVTPGSRCPRRCACYVPTEVHCTFRDLTSIPDGPANVER
 VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQLQVLKMSYNKVQII
 KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLRLVHLEGNRLTKLHPDTFVSLSYLQIF
 KTSFIKXLYLYDNFTSLPKEMVSSMPNLESLYLHGNPWTCDCHLKWLSEWMQGNP
 (SEQ ID NO: 15)

Figure 18

MKVKGRRGITCLLVSVFAVICLVATPGGKACPRRCACYMPTEVHCTFRYLT SIPDSIPP
 RINLGYNLSVRLMETDFSGLTKELEMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
 LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYL
 QIFKISFIKFLYLSDNFLTSLPQEMS YMPDLSLLYLGPNWTCDCHLRWLSDWIQPDVI
 KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
 AFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYTVLNTSFSTFL
 VCNIDYGHIQPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
 ADPSWLMQDQISLQLNRTATTFSTLQIYSSDAQITLPRAEMRPVHKW
 TWMISRDNNTK
 LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRA
 PYVSEDGRILIDKSGKLELOQM
 ADSFDTGVYHC
 ISNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPC
 HSTGIP
 DASISWVIPGNVLYQSSRDKKVLNNGTLRILQVTPKDQGYYRCVA
 ANPSGVDFLIFQV
 SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTS
 ALMEAEGKHTSSTS
 KRH
 NYRELT
 LQRRGDSTHRRFRENRRHFPPSARRIDPQHW
 AALLEKAKKNAMPDKRENTTV
 SPPP
 VVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTAD
 SRTISD
 SPMTNIN
 YGTEFSPVVNSQILPPE
 EPTDFKLSTAILTTAMSKNINPTMSSQI
 QGTTNQHS
 STVFPL
 LLG
 ATEFQDSDQMGRG
 REHFQSRP
 PITV
 RTMI
 DV
 NV
 KML
 S
 TT
 N
 K
 L
 L
 E
 S
 V
 N
 S
 H
 Q
 T
 S
 V
 R
 E
 V
 SEPRHNHFY
 SHTTQIL
 ST
 T
 F
 P
 S
 D
 P
 H
 T
 A
 A
 H
 S
 Q
 F
 P
 I
 R
 N
 S
 T
 V
 N
 I
 P
 L
 F
 R
 R
 F
 G
 R
 Q
 R
 K
 I
 G
 G
 R
 R
 I
 I
 S
 PYRTPV
 LRRH
 RYS
 IFR
 S
 T
 TRG
 S
 SEK
 S
 T
 A
 F
 S
 A
 T
 V
 L
 N
 V
 T
 C
 L
 S
 C
 L
 P
 R
 E
 L
 T
 T
 A
 T
 A
 A
 L
 S
 F
 P
 S
 A
 A
 I
 T
 F
 P
 K
 A
 D
 I
 A
 R
 V
 P
 S
 E
 E
 S
 T
 T
 L
 V
 Q
 N
 P
 L
 L
 L
 E
 N
 K
 P
 S
 V
 E
 K
 T
 T
 P
 T
 I
 K
 Y
 F
 R
 T
 E
 I
 S
 Q
 V
 T
 P
 T
 G
 A
 V
 M
 T
 Y
 A
 P
 T
 S
 I
 P
 M
 E
 K
 T
 H
 K
 V
 N
 A
 S
 Y
 P
 R
 V
 S
 S
 T
 N
 E
 A
 K
 R
 D
 S
 V
 I
 T
 S
 S
 L
 G
 A
 I
 T
 K
 P
 P
 M
 T
 I
 I
 A
 I
 T
 R
 F
 S
 R
 R
 K
 I
 P
 W
 Q
 Q
 N
 F
 V
 N
 N
 H
 N
 P
 K
 G
 R
 L
 R
 N
 Q
 H
 K
 V
 S
 L
 Q
 K
 S
 T
 A
 V
 M
 L
 P
 K
 T
 S
 P
 A
 L
 P
 Q
 R
 Q
 S
 S
 P
 F
 H
 F
 R
 R
 L
 S
 T
 S
 V
 M
 Q
 I
 P
 S
 N
 T
 L
 T
 I
 A
 H
 H
 I
 T
 T
 K
 T
 H
 N
 P
 G
 S
 L
 P
 T
 K
 K
 E
 L
 P
 F
 P
 P
 L
 N
 P
 M
 L
 P
 S
 I
 I
 S
 K
 D
 S
 S
 T
 K
 S
 I
 I
 S
 T
 Q
 T
 A
 I
 P
 A
 T
 T
 P
 T
 F
 P
 A
 S
 V
 I
 T
 Y
 E
 T
 Q
 T
 E
 R
 S
 R
 A
 Q
 T
 I
 Q
 R
 E
 Q
 E
 P
 Q
 K
 K
 N
 R
 T
 D
 P
 N
 I
 S
 P
 D
 Q
 S
 S
 G
 F
 I
 T
 P
 T
 A
 M
 T
 P
 P
 A
 L
 F
 T
 H
 S
 P
 P
 E
 N
 T
 G
 I
 S
 S
 T
 I
 S
 F
 H
 S
 R
 T
 L
 N
 L
 D
 V
 I
 E
 E
 L
 A
 Q
 A
 S
 T
 Q
 T
 L
 K
 S
 T
 I
 A
 S
 E
 T
 T
 L
 S
 S
 K
 S
 H
 Q
 S
 T
 T
 R
 K
 A
 S
 L
 D
 T
 P
 I
 P
 P
 F
 L
 S
 S
 S
 A
 T
 L
 M
 P
 V
 I
 S
 P
 P
 F
 T
 Q
 R
 A
 V
 T
 D
 R
 G
 D
 S
 H
 F
 R
 L
 M
 T
 N
 T
 V
 V
 K
 L
 H
 E
 S
 S
 R
 H
 N
 L
 Q
 M
 P
 S
 S
 Q
 L
 E
 P
 L
 T
 S
 S
 N
 L
 L
 H
 S
 T
 P
 M
 A
 L
 T
 T
 V
 K
 S
 Q
 N
 S
 K
 L
 T
 P
 S
 P
 W
 A
 E
 Q
 F
 W
 H
 K
 P
 Y
 S
 D
 I
 A
 E
 G
 K
 K
 P
 E
 V
 S
 M
 L
 A
 T
 T
 G
 L
 S
 E
 A
 T
 T
 L
 V
 S
 D
 W
 D
 G
 Q
 K
 N
 T
 K
 K
 S
 D
 F
 D
 K
 K
 P
 V
 Q
 E
 A
 T
 T
 S
 K
 L
 L
 P
 F
 D
 L
 S
 R
 Y
 I
 F
 E
 K
 P
 R
 I
 V
 G
 G
 K
 A
 A
 S
 F
 T
 I
 P
 A
 N
 S
 D
 A
 F
 L
 P
 C
 A
 V
 G
 N
 P
 L
 P
 T
 I
 H
 W
 T
 R
 V
 S
 G
 L
 D
 L
 S
 R
 G
 N
 Q
 N
 S
 R
 V
 Q
 V
 L
 P
 N
 G
 T
 L
 S
 I
 Q
 R
 V
 E
 I
 Q
 D
 R
 F
 Q
 Y
 L
 C
 S
 A
 N
 L
 F
 G
 T
 D
 H
 L
 H
 V
 T
 L
 S
 V
 V
 S
 Y
 P
 P
 R
 I
 L
 E
 R
 R
 T
 K
 E
 I
 T
 V
 H
 S
 G
 S
 T
 V
 E
 L
 K
 C
 R
 A
 E
 G
 R
 P
 S
 P
 T
 V
 T
 W
 I
 L
 A
 N
 Q
 T
 V
 V
 S
 E
 S
 Q
 G
 S
 R
 Q
 A
 V
 V
 T
 V
 D
 G
 T
 L
 V
 L
 H
 N
 L
 S
 I
 Y
 D
 R
 G
 F
 Y
 K
 C
 V
 A
 S
 N
 P
 G
 G
 Q
 D
 S
 L
 L
 V
 K
 I
 Q
 V
 I
 A
 A
 P
 P
 V
 I
 L
 E
 Q
 R
 R
 Q
 V
 I
 V
 G
 T
 W
 G
 E
 S
 L
 K
 L
 P
 C
 T
 A
 K
 G
 T
 P
 Q
 P
 S
 V
 Y
 W
 V
 L
 S
 D
 G
 T
 E
 V
 K
 P
 L
 Q
 F
 T
 N
 S
 K
 L
 F
 L
 F
 S
 N
 G
 T
 L
 Y
 I
 R
 N
 L
 A
 S

SDRFTYECIATSSTGSERRVVMLTMEERVTSPIEAAASQKRTEVNFGDKLLLNCATGEP
 KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGYLCVARNKGDDLILMHVSLR
 LKPAKIDHKQYFRKQVLHGKDFQVDCASGSPVPEISWSPDGTMINNAMQADDGHR
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMVKHLTVITAAPRIRQSNKTN
 KRIKAGDTAAVLDCEVTGDPKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
 EYVCVARNPMSGDDTKMYKLDVSKPPLINGLYTNRTVIKATAVRHSKKHFDCRAEGTP
 SPEVMWIMPNDNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGESVLVVQLE
 VLEMLRRPTFRNPRNPNEKIVAQLGKSTALNCVSDGNPPPEIWLPNGTRFSNGPQSY
 QYLIASNGSFIIISKTTREDAGKYRCAARNKVGYIEKLVILEIGQKPVILTYAPGTVKGISGE
 SLSLHCVSDGIPKPNIKWTMPSGYVVDRPQINGKYILHDNGTLVIKEATAYDRGNYICKA
 QNSVGHTLITVPVMIVAYPPRITNRPPRSIVTRTGAAFQLHCVALGVPKPEITWEMPDHS
 LLSTASKERTHGSEQLHLQGTLVIQNPQTSIDSGIYKCTAKNPLGSDYAAATYIQVI (SEQ ID
 NO: 16)

Figure 19

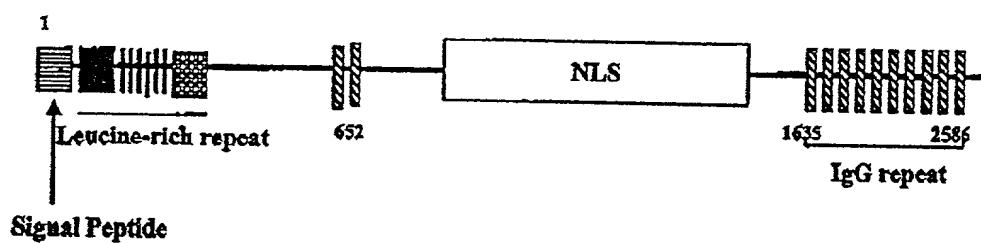


Figure 20

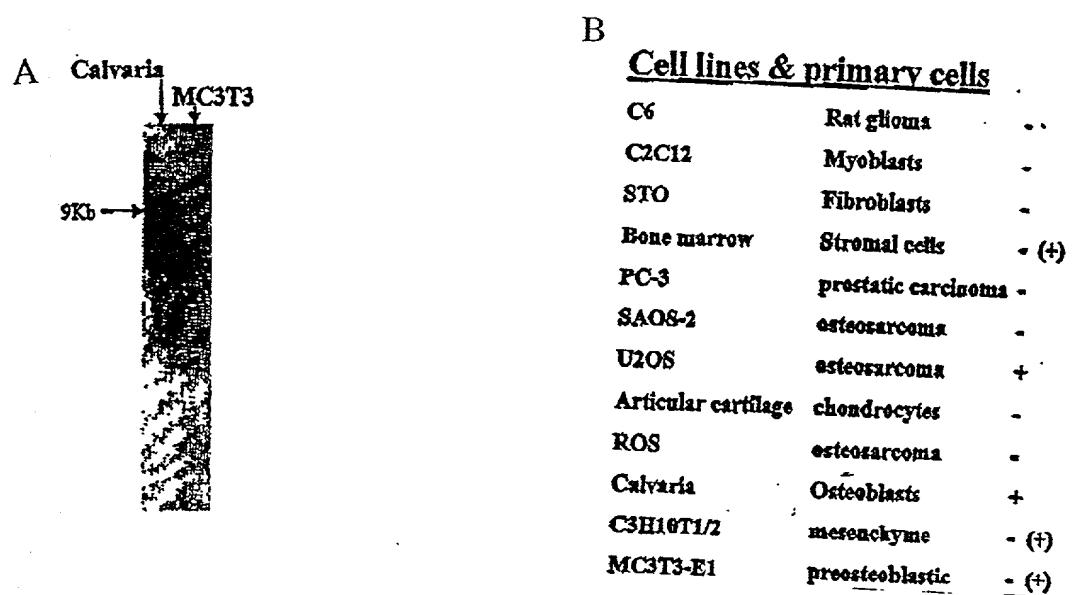
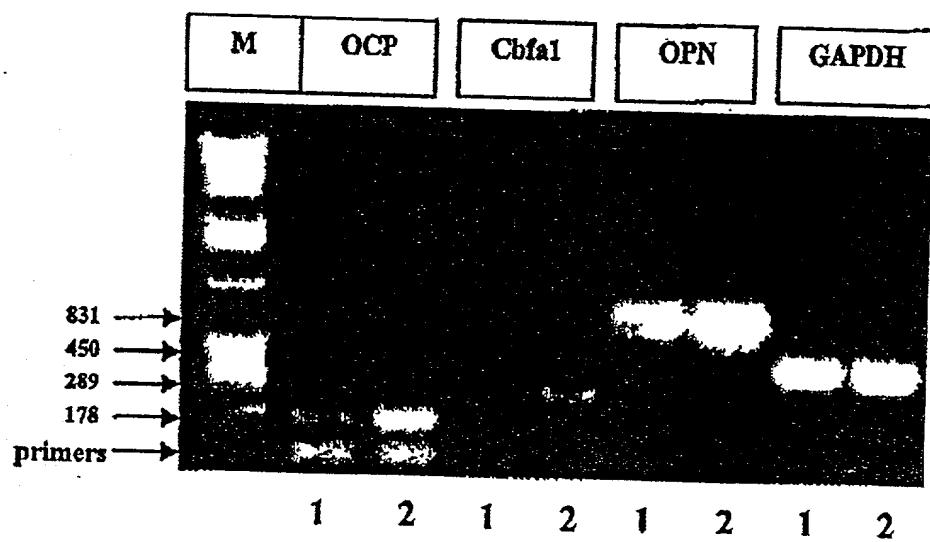


Figure 21



51/90

Figure 22

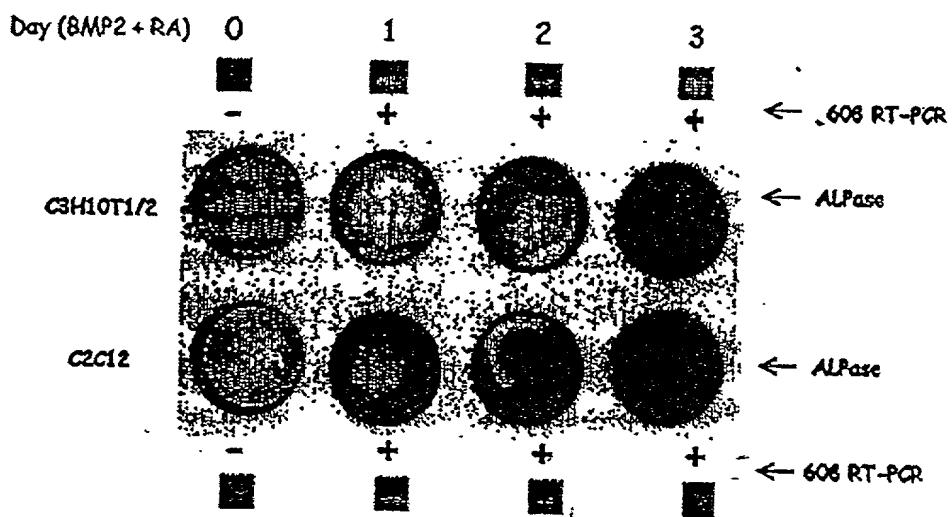
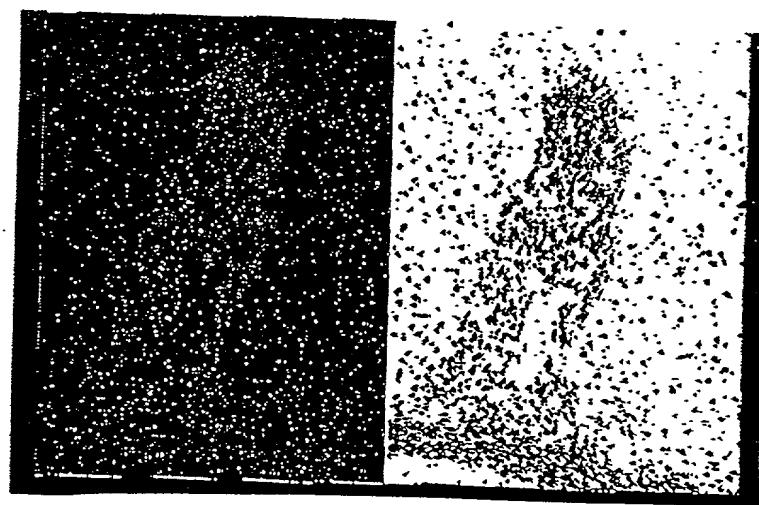


Figure 23



52/90

Figure 24

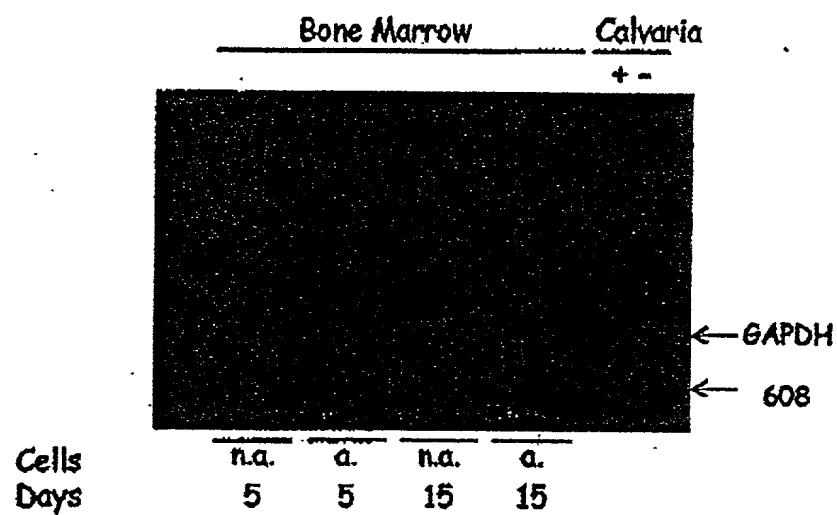
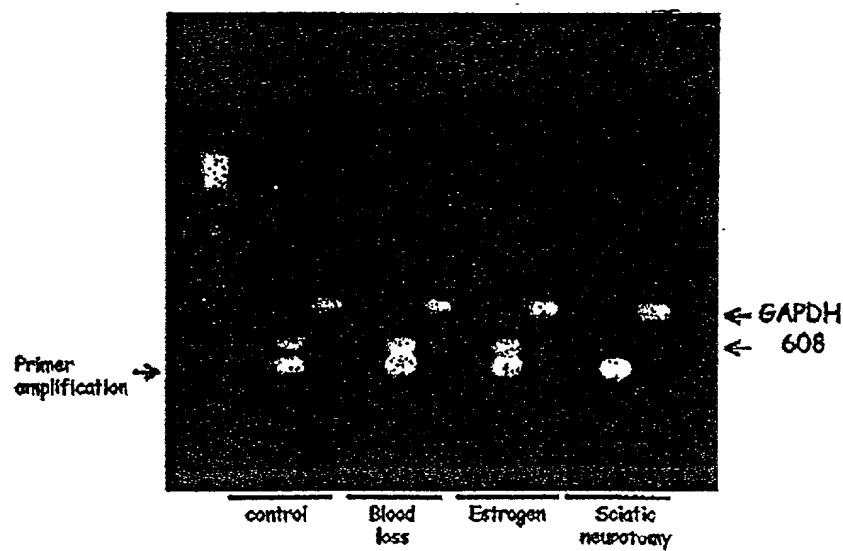


Figure 25



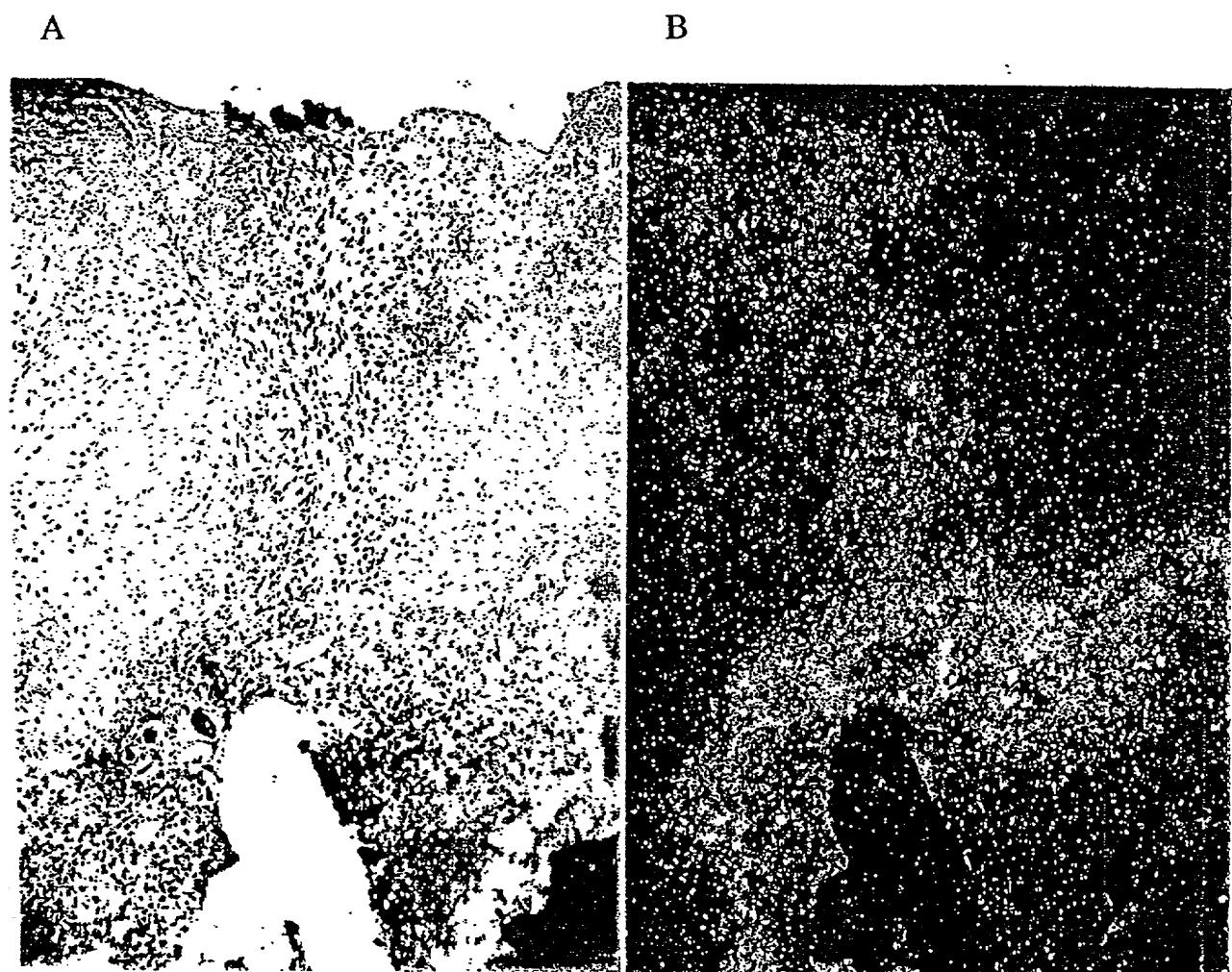
53/90

Figure 26



54/90

Figure 27

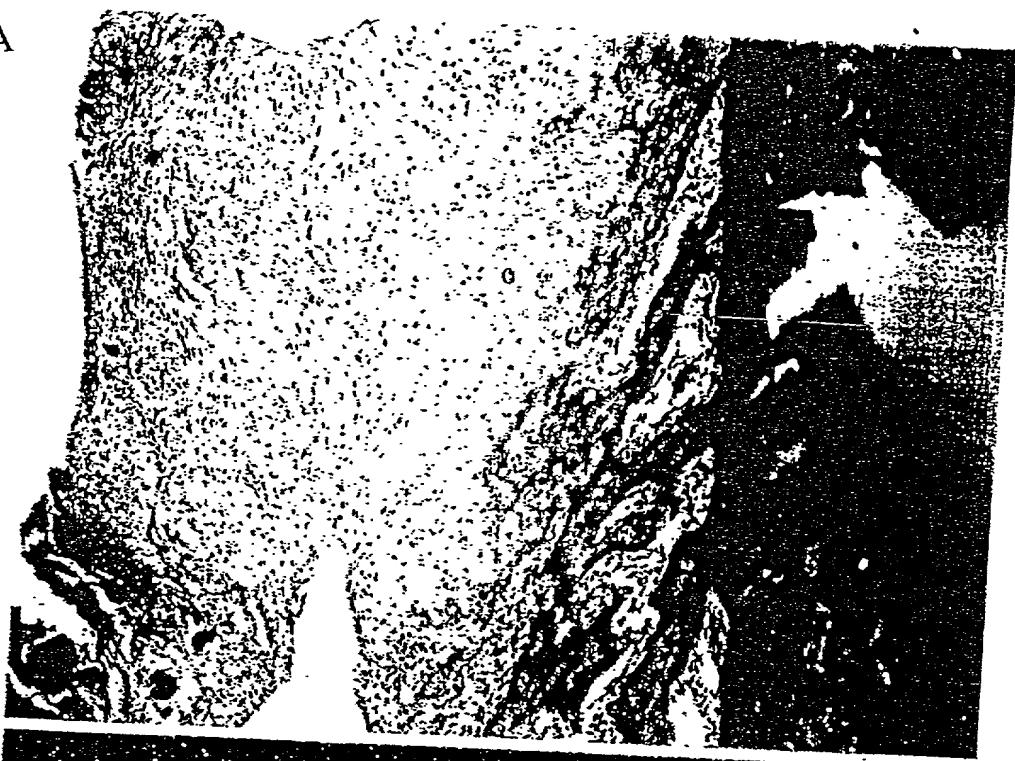


2020-06-06 09:50:54

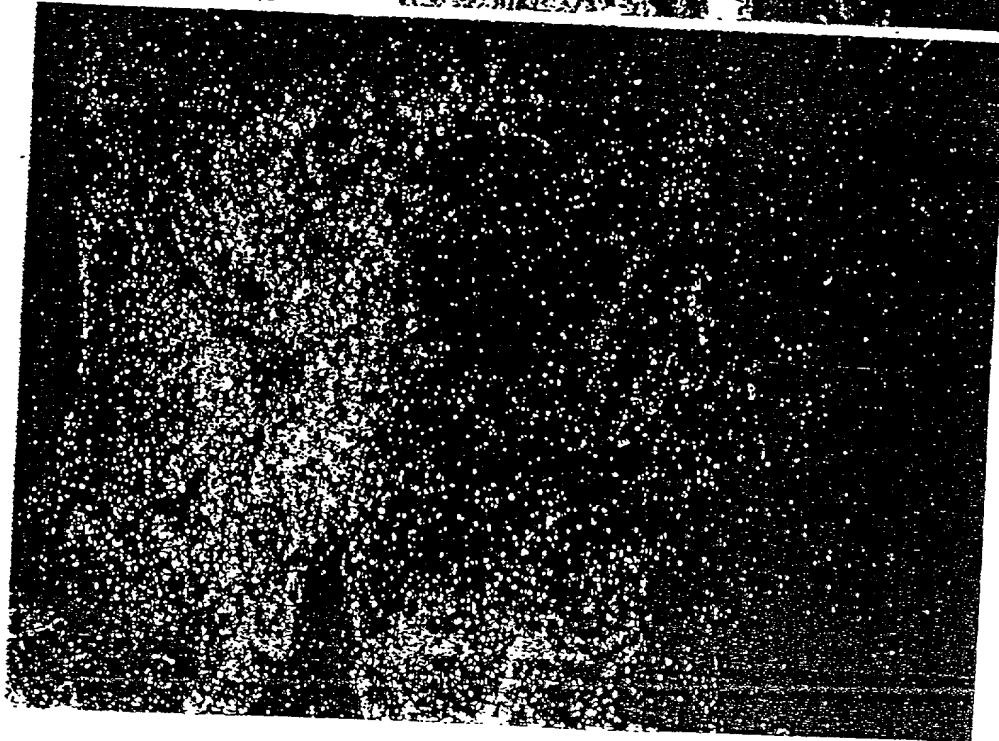
55/90

Figure 28

A



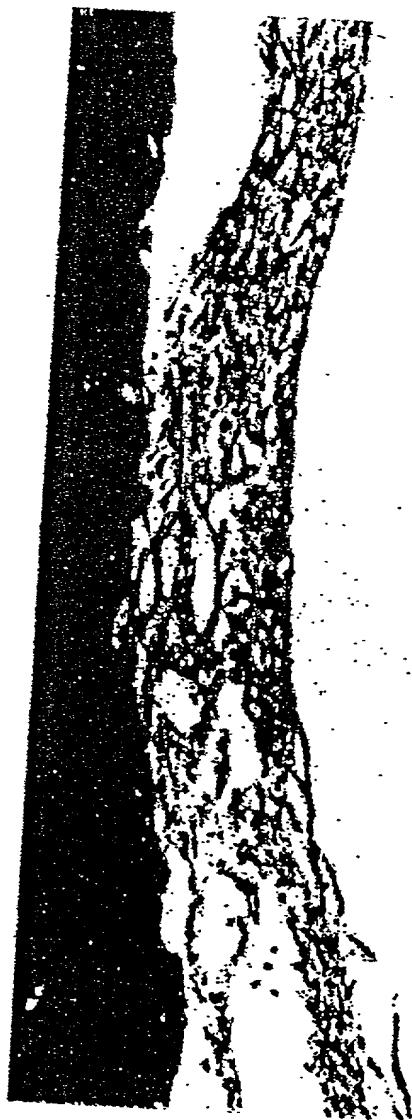
B



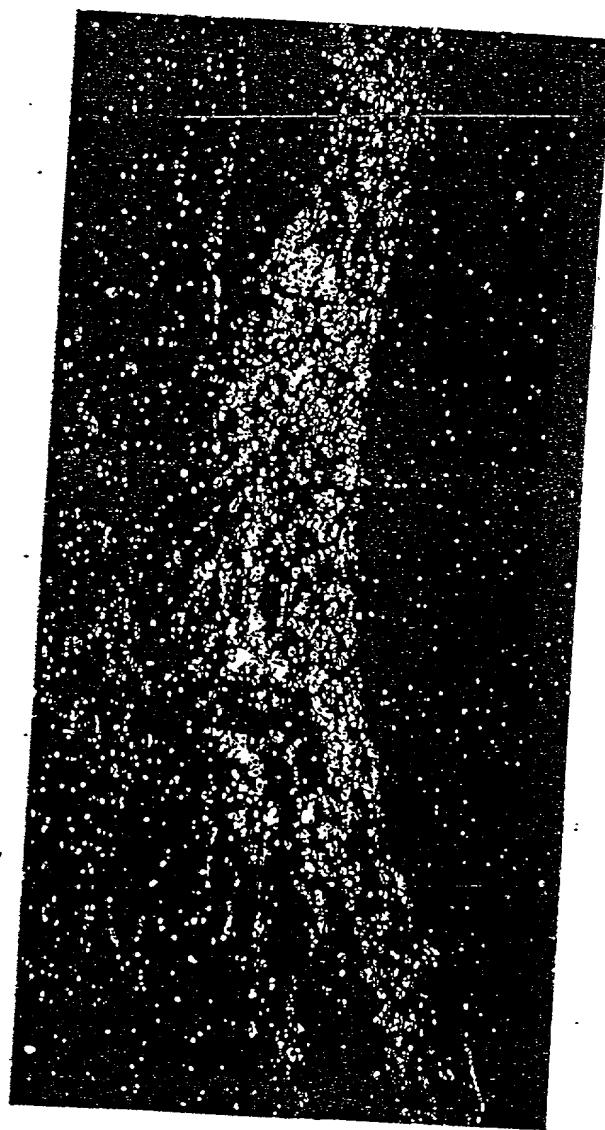
56/90

Figure 29

A

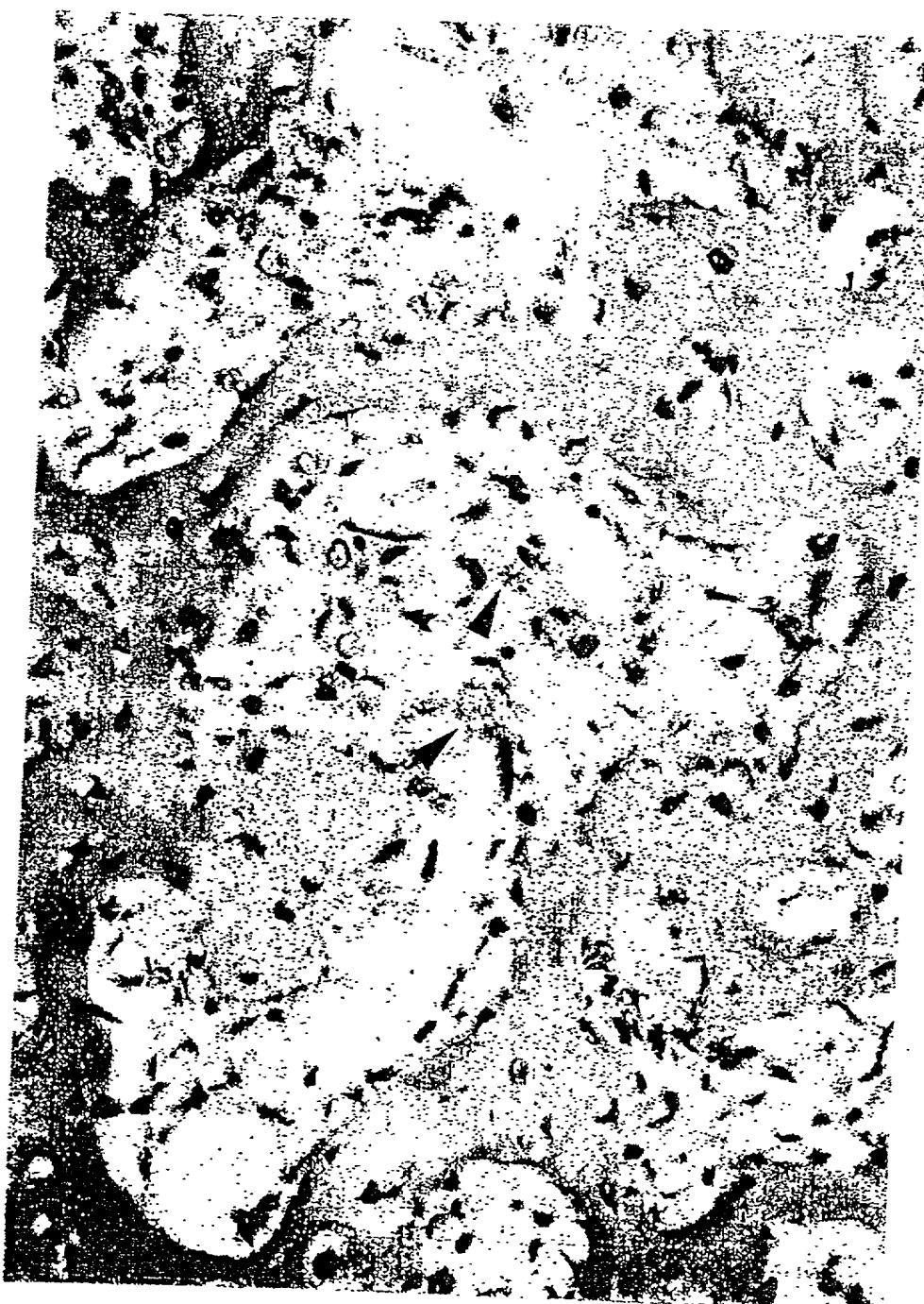


B



57/90

Figure 30



58/90

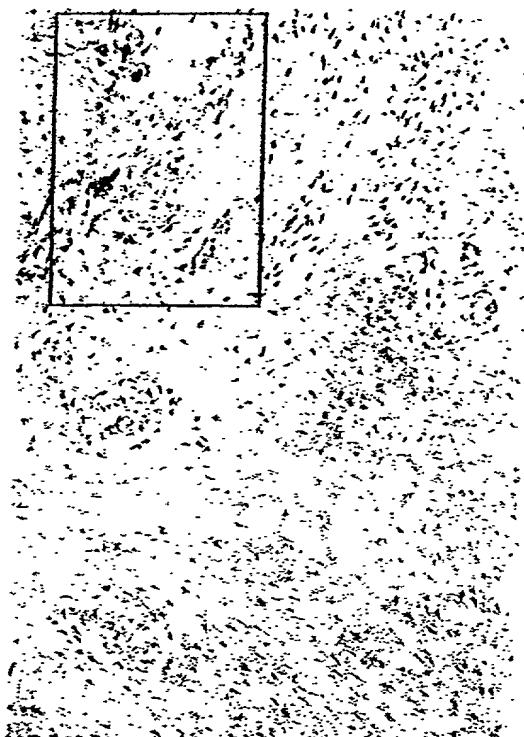
Figure 31



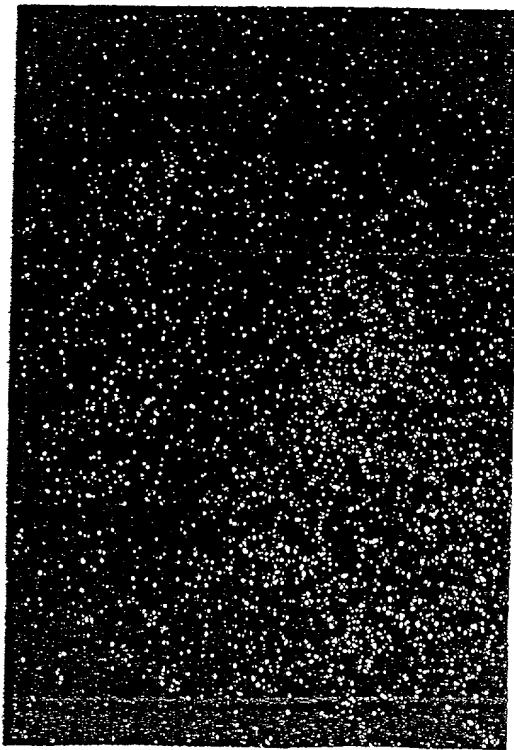
2022-07-20 10:56:00

Figure 32

A

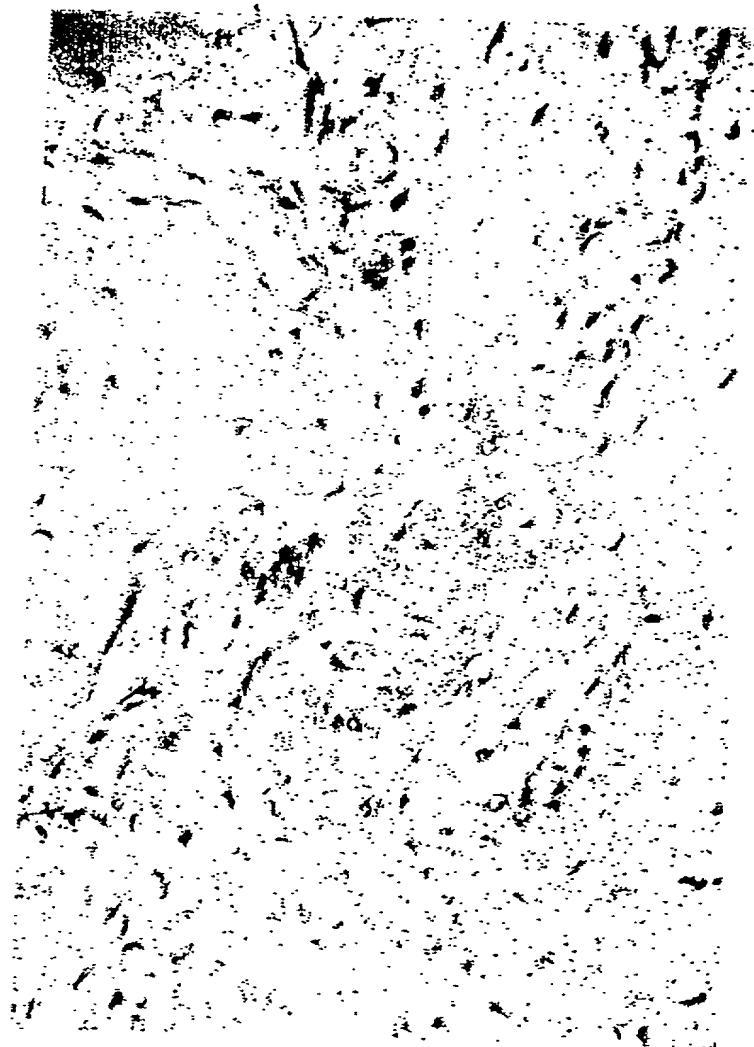


B



60/90

Figure 33



090909-0202-022202

61/90

Figure 34

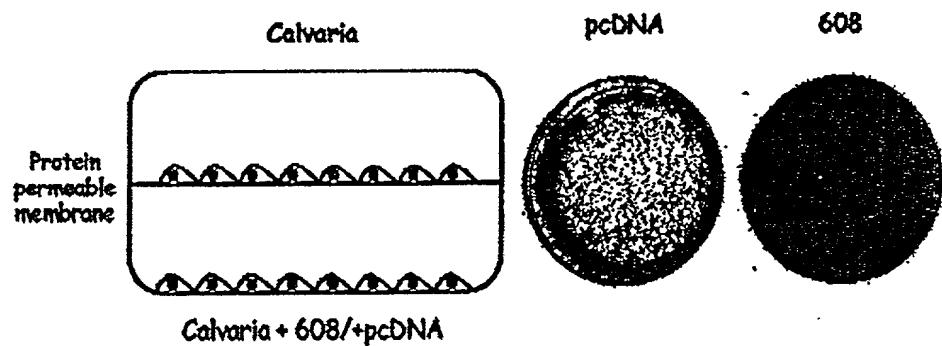
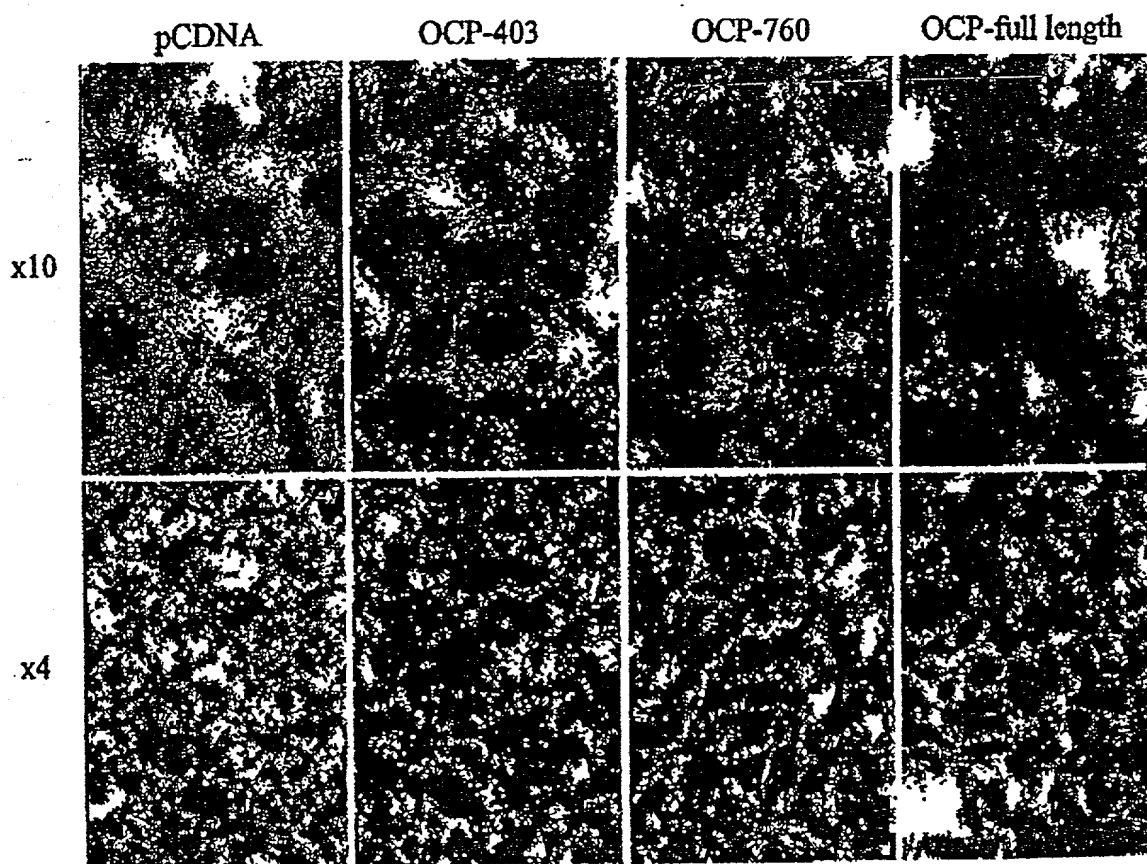


Figure 35



62/90

Figure 36

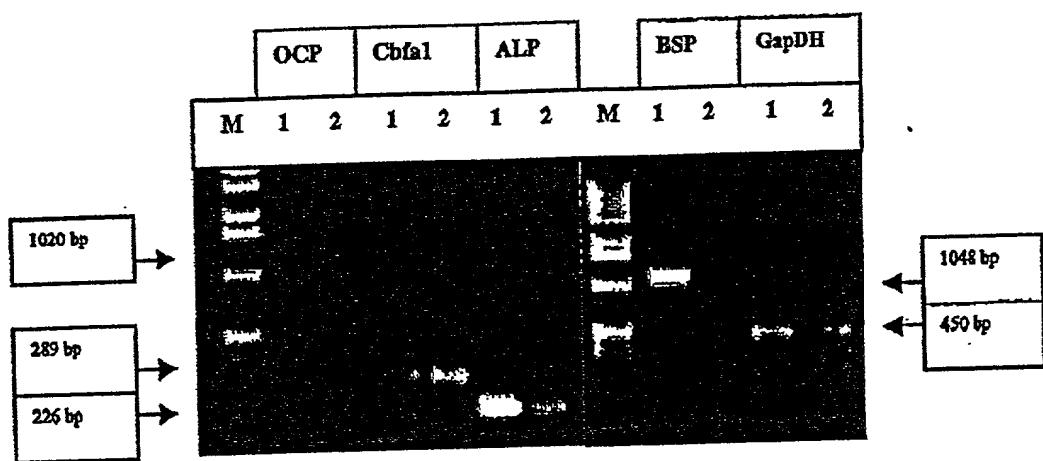


Figure 37

x4

x10

pCDNA
ROS stable line

OCP
ROS stable line

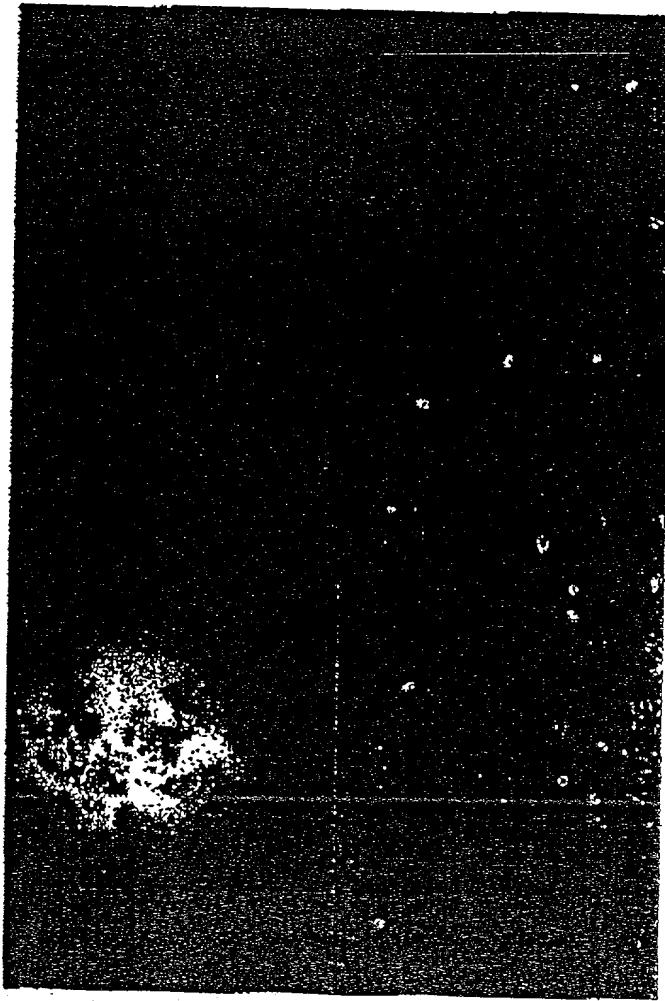


Figure 38

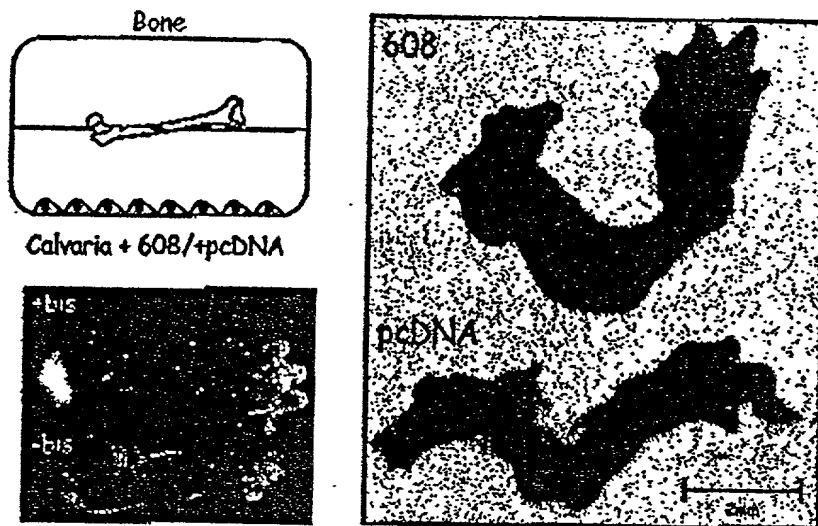
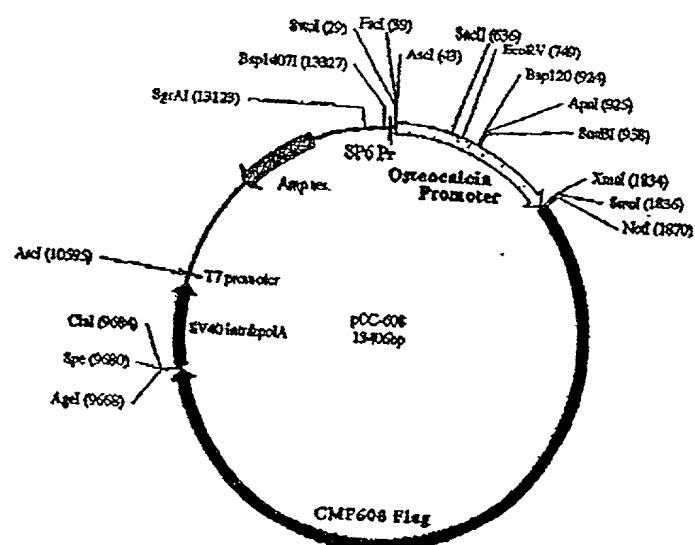
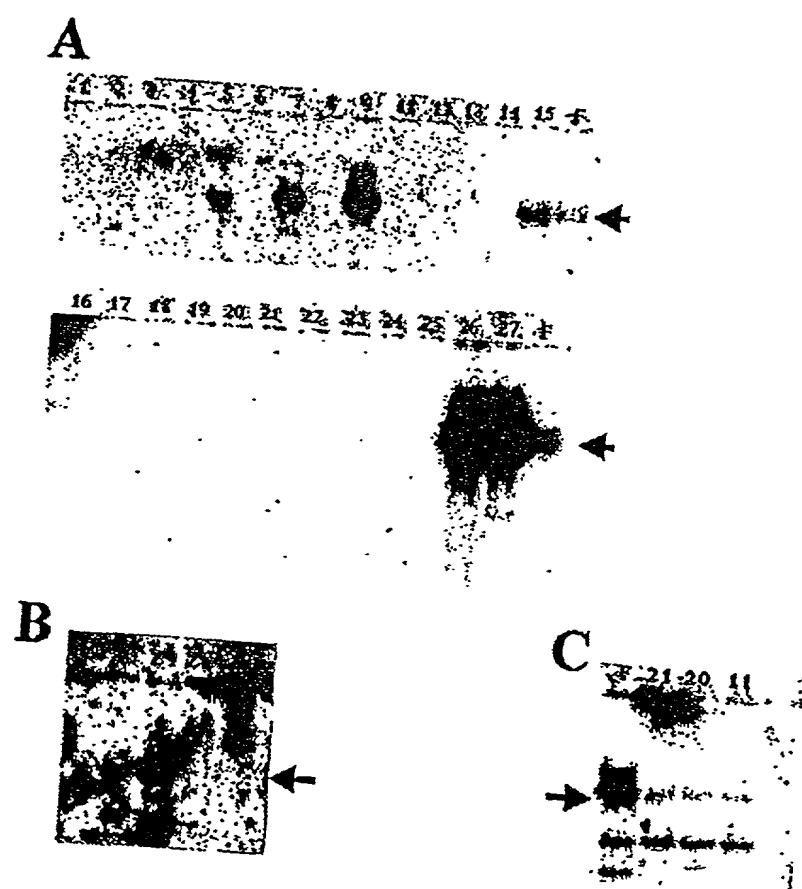


Figure 39



64/90

Figure 40

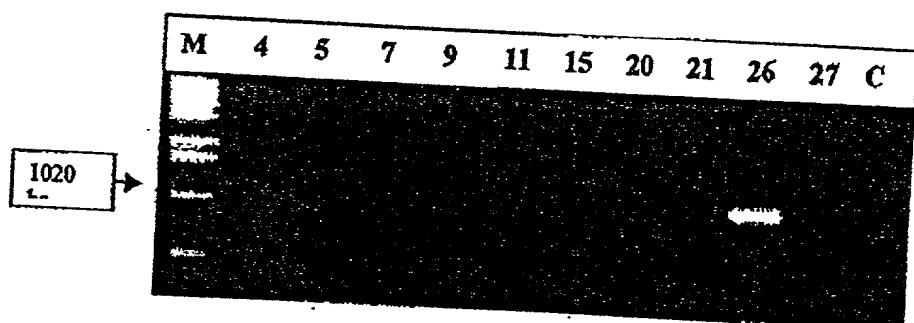


000000000000000000000000

65/90

Figure 41

A



B

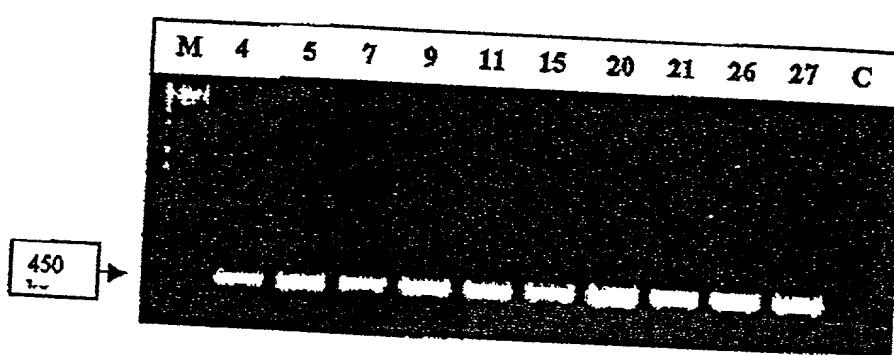
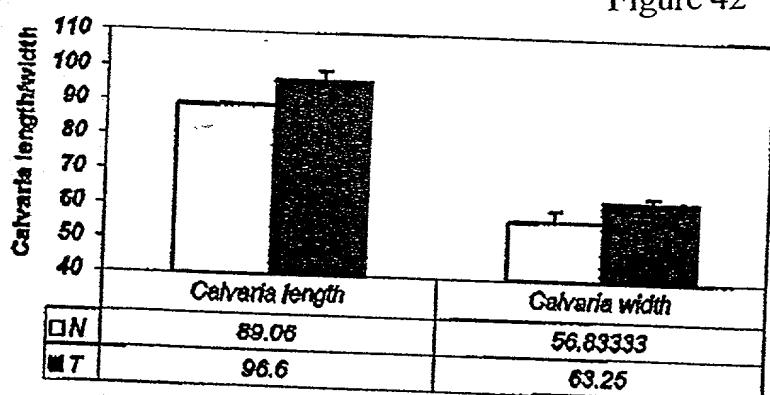
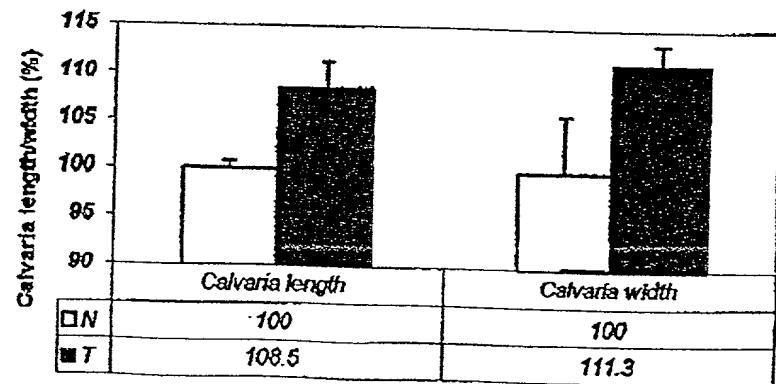


Figure 42

A



B



66/90

Figure 43

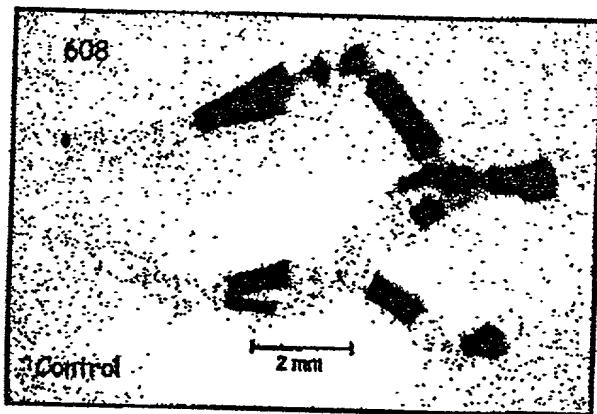
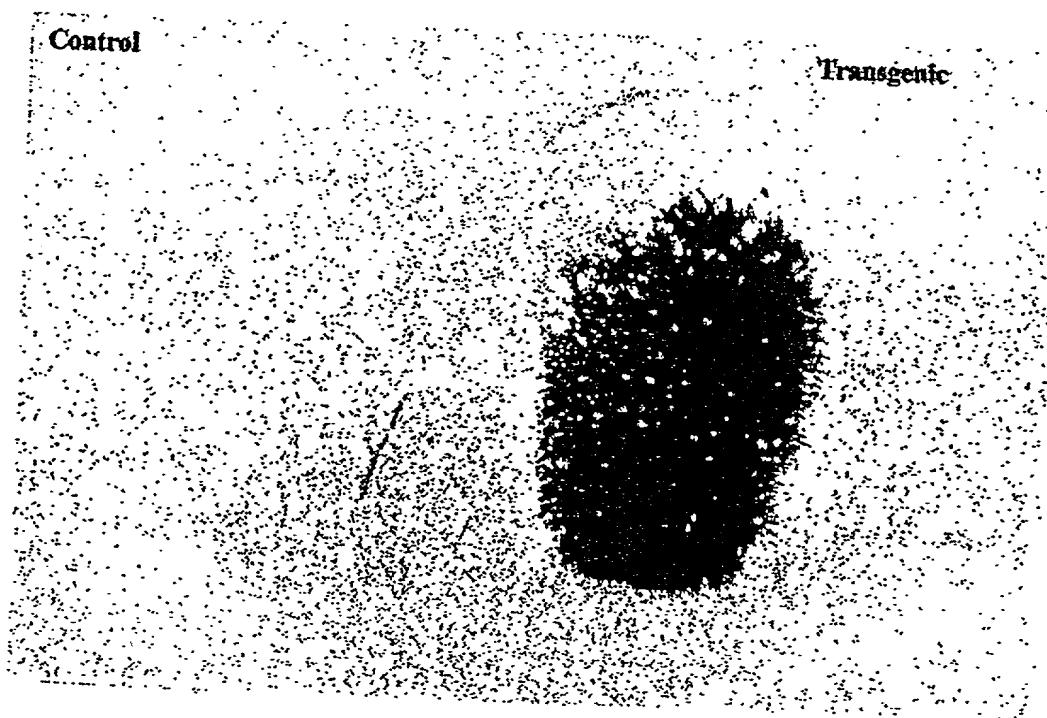
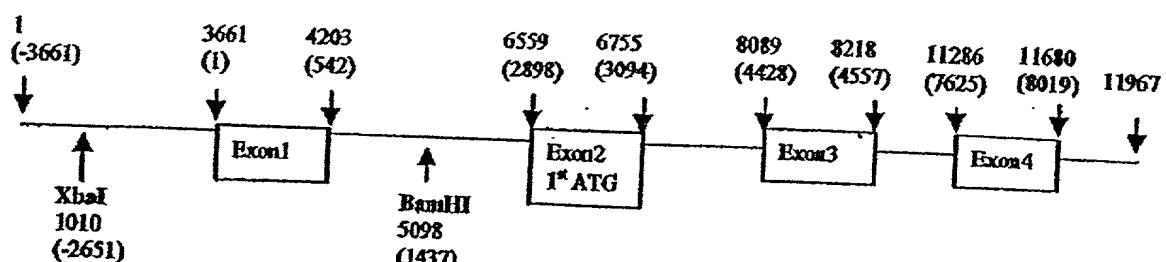


Figure 44



67/90

Figure 45



Clone 14C10 / 1SE11

← T7 orientation in pKS

Figure 46

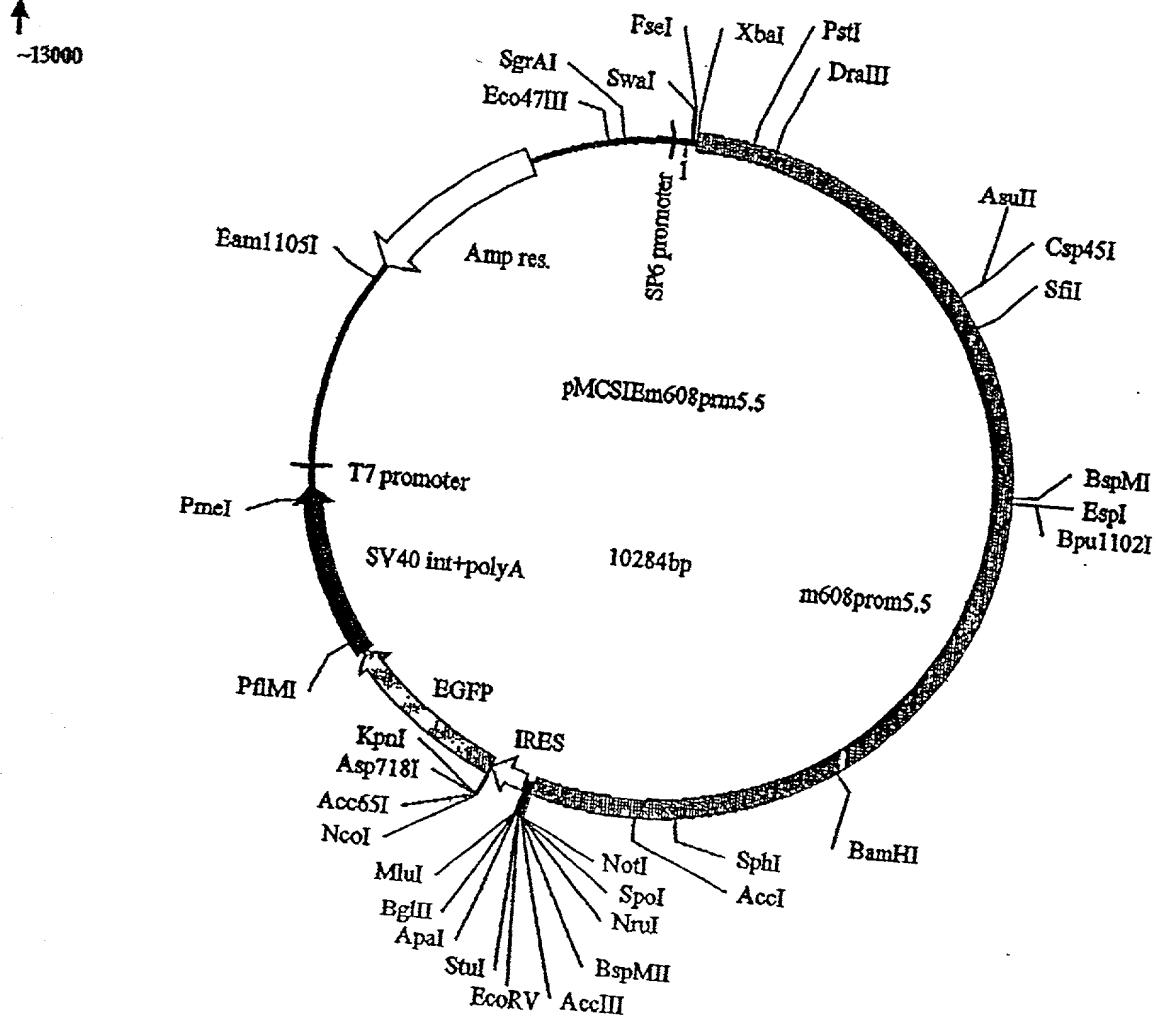


Figure 47

↓ (XbaI)
 TCTAGAAGTAAATGATCTGAGTAGCGATCTGGAAAATACGTACTCTAACACACTGC
 AATCATCTCTGTGGTTGCTGGAGCTGAGGTCTGGAAGGCTGACCTTGGTTAGAAAATA
 ACCTACCGAATACAGAGCTATGACGTTAGCTGGAGGAGCTTGGAAAGAATGACAAGCT
 GTAGCTGCCAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCCGACAT
 CTTAAGAATATGGTGTCTGAGTGTCAAGGACCTTGTCTGCCACACAGGTCTCCAG
 GTTATGTGCTAACTCTGACTGCTTAAATTCTACCTTGTATCATTAAATGACTAGAA
 ATCACTTGGTATTAGCAACTGGATATGGAATATTACTAATTGTACCCAAGGCCAGGCC
 CTCAGCTTGGCAGCTCCATTCTCTGTGGAGGCCAGTCACGTGGGTTGAATCAACTGT
 ACTGTCTACTTACAAGACGCAATTACCTGAGATGAGTCATTCTCACAAGCTTTTA
 GAAGAGTCATTAGACATATTCTGATGAAGTAAGCATATAAAAGTGAGAGCAGCATGAATG
 TGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAATAAACTGACTTTAAAAAAA
 AGCTTATGATACTTGTACAGAGTAAATCTCCTAAATATCATCTGCATTATAAATAT
 TTCTATAATCCATCAATTAAAAACCTTAAAGAAATTGTAAACCAAAGATCCCTAGGCC
 CTGCCCCCTAGGATGGCTGTATGGTGGGCCCTGAGAGATGGCTTAAAGAACTTACTTGTCC
 AGGAGCACATCTCAGAACATCTGCCCTAAACATTATCCAAATGCTCATCAAAGGCTC
 ACTCACATGTGCTICAACCAAGGGATTAAACAGTCATTAGTCACATTCTCACAAGGT
 GGAAGCTGCTAGAGGAACAGGGATGTACAGGATAACATCCAACCTTACAAAAGGATGTC
 ATAACCCCTACCACAAACAAACAACAGACAACAAACCCATAAAATTATCACGGCAAAT
 GAACTAAGCCATATGCAAGAAAAGTATTATATGTTCTCATTTGGGGTGTCTTCTTAAT
 AGTCAAATATGCAAGAATATGACAAAGATGGTTATGCAAGTGGGATGGCGAAGGATA
 CTTGTAGATTAGAGGACACAAAGCAACAACTACAGAGTGAAGTAATCCAGAGACTTAATG
 TATAATATGAGGACTGTATTAAATAATTCTATTAAAGATAACACAGCAAACGAGTGTATCTT
 ACTAACACACACACTACATAGAGAGAATAAAAGTGATAGATACTGTTGTCTTATCTCAG
 TAGCTGATAATTCTATTTGTACACCTCAACATAGATAACCAACAAAGAGGAAGAGGAT
 AGGTGCTCTCCCAGGGCGGAAGAGTACATTGCAAGTCAGACACCATTGTGAGATGTA
 CCACATGGAGGAGCTAGAGAAAGTAGCCAAAGGAGCTAAAGGGATCTGCAACCCCTATAAG
 TGGAAACAACATTATGAGCTAACAGTACCCCCGGAGCTCTGACTCTAGCTGCTATATATC
 AAAAGATGCCCCTATGGGCAATTCTGAAAGAGAGGGCCATTGGACTTGTCAAACCTTAT
 ATGCCCCAGTACAGGGAAATTACCAAGGGCCTAAAGGGGAGTGGGTGGGAGGGAGTG
 GGGGTGGGTGGATATGGGGACTTGTGTTAGCATTGGAAATGAGTTAAATAC
 CTAATAAAAATGGAAAAAAGGAAAGGAGGTCAAGCACCTCAC
 TTCACTGCTATCTCAACTTGTAAACAGAACAGGGAGTCACAAACCCAGGACAAACCAAGT
 GATTGAAGCGTCTTGTAAATGTATTGCTGTTGTTACCACTCATTAGCATATATTCT
 TGTGAAAACCTACGGGGCTATGACATGTTTTTATTCAAGTATATCACATGCTGTCAGC
 ATATTGGCACCACACTACCAAGCCCCAGCCCCCTTGTCCCCGCCCCAACACACACACACA
 CAC
 CATCATCTGCTCACTCACCCACCCAAAGCTTAATCTTCTTCTGCAATAGTACCTCTC
 CTATTTTATGTCTAGGTTCCCCCTCCCCCTGTAGGAGATGGGAGAGGTACGAAAGGAA
 AGAATTGTAGCCCCCTGAGCCAGCCGGGCCACAGAGCCTGOCACCCAGACAGGAAAGCC
 CAGGGCTTACCAAGCACAGGAGGAGCAACTCGCAGGAGCTGGGTGGCGCTGGTGGT
 CCCGGTCTGATGGCCCGCCATTCCAGAAGCCGAGGCTATAGCTGCGTCACCTGCCCCG
 CCCTCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGACCCAAAGGCGGTGCAATTCCA
 TTGGCCCAAGGCAGAGGTGAGCGCTGCTAATCCCTGGAAAGTGAAGGGACCCAGAG

AGATTTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAAGGTTATGCCAACCTCTTGTAAATACAGTTGAGGCTTAAGTGTACGGGAACCTCATGTGGTA
TTCATTTCACGGCTCTCTCTTATAACTAACTCTTAAGGTGCTATAGTCTCTCTGTTCTCCAGCTACCTTGACCATCTTGTATCTAATAATAGCAAGCTCATCTGTTTAAATCATC
ACGCAGAGAGTATTCAAAAATATTCACTGATGTAACAGTGTACAGTGTAGGCATAGAAGTA
ATCATTAGTAAATCTTAAATTGGGTTAAACTCATTCATAACAGCTCCAGGTGGGAGGGAT
CACTGAGCCTTGGCCACAGTGCGGGGTTAAAGATATTCTAACAAGAGAACAGAGAATTCTC
CTTGGCCATGCTCCCCATCACTGTGTCACTAAGCAGAGGGGTGTTCCAAGCAGAGAAAG
AGCAGACAGTGTATGCTGCAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTACT
GTCTCCCGGTCTTCTAGTGTGCTCTGAAAAGGCCATGTGTCTTATGGCAAGGACTTGC
AGACATGCTAGAAAGAAATTGACCTTTCTAGTGGGTATTACAGCTGTTAAAGTAT
TTGGAAAGGTAAAGCCAAATAAAATACACATATTAAATAATACAATGTTACAAAAAATT
GATCATATAAGAAGTACATTCTAAATGCAATGTGAAAAATATATATAATTCTATCTAT
TTACTGGTGCCTAAAGTTCTAAATTGACATGTACTATTCTATATTATAAAATATTCT
AAAATGTATATAAAAGTGTAAAGGCTCTGGTCAAAACAAGAGAGTTAAATTACAAACT

↓ (BamHI)

TTAATTGTCCCCATAACATTATTATGATCTCTAATGACAGGGATCTCTGCTTTCATGGGA
 AATGAGAAAGCTATGAAGATATGTTACAATAATAAGCCCCATTAGTGTATAAGTCCAATG
 GGAAGCTAGCACACACTGGTTATAAAGAGAACAGTTCTGAGTCTATGCAAGTTACA
 CTCTAGGGAATAAGAGTTCCTCTTCTCAGATTCACTAGCATTGTTGTCATCATTATC
 TTCTTGATGATGAGCATTATAAGTGGATAAAGATAGGATCTCAAAGGAATGTCATTG
 ATGCCCTGAACAATCTTCAGGCTTCTCAGTTCACTAGTCTATTCAATTATGGATAA
 TTGGGGGATGGTGTAAATTTCAGTTATGAAATTCCAAAAACAAAAACAAAAACAAAC
 AAACAAACAAAAACCTCTGAAACTAGAACTACCAATCATTACTGGGTATGTAACAAAG
 AGAAATCTGCACAGAATTATGCTACATGTCATTACGACAGCCAAGAATGTTGGA
 ACCAACATACGTAGCCGTCAAAATATGAAACGGATAAAGAAAATGTTGAAATGTTGTAACAC
 AGAGTCCCATGTGGCCATAAAAGAGTGAATCATGACATATGCAGGAAATGGATGCAACT
 GGAATCAATTGGCTAATCAAAACAGACAGACTCAAAAGGAAACACCGTGTAGCT
 CTCTGACAAAACAGAAAGCTAGATTACACTTGTACGGTCCATGTTGTTGTTAGAATTATT
 TAGTTATACACTATTCTAATCTGTAGTGTATAAAGGATGCTATGAAAGCAAAACA
 AGCTAGCTGGGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTATAAGAACAGAACATA
 GTAACATAGGTGCCAGGATGAAATGCTTAAATTGTCTAATCCTAAACACAGACAGGA
 GGCACACGTTCAACCCAGGGTGAATCCTCAGCACAGAGAAGGGGAAGTAGACACAAAGT
 TTGCCACTAACCAAGAACGCCATTGCACTTGTGCTGCTGCTGGAGGGCGTTCCAGTTT
 CTCCAGTCTGACACTGTGTATAACAAACAGTGTACAATACAAAGTTGGCATGATGGATGG
 TTTTGTGTATTTCATTTCATTTCATTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
 GCTGTGGTTTCATTGTTCTTGTGAGAGAGAGAAGGAACATGAAATTGGGTGGGTAGGA
 AGCTGGAAACGATCTGGAAGAAGTGGGGAAAGAGAAAAATTGTATGGAGCATAATTAA
 AAAAAACAAACAAACAAACAAAGGTCTATTGCAACAAAAGGTGTGAATTAAATTAAAC
 CAGTTACGACTCTAAAGAAAATATCCCAATTATCCCAAGAGTGTCTATGATGCTGTG
 CTTAGGACTTTGCTGAACTGGCCCTATAACTCTGGTGTGGTGTCTTTCAGG - Artificial

Notable

(SEQ ID NO:17)

Figure 48

CACAGACCTTCCCTTCTAACCTCTCCTCCCCATCTTGTGCTTCATCCAGACTTCACAC
 CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCTGTCAGGACAACAAGGAGGCTGAAG
 GCAGACCCACACGTTCCAACCTGCTCTGAGAGTCATCCCCCTAGACTCATCTATAAGCAG
 GAAACCTGCTGTGATCTCCATTCTCTGACCACATCCCCAAGTTATCACAAGGAGTT
 TTCTCTCAAAACCTTCTCTCCAGCAAACCCCTCACTGCTCTGGGTACTTTCTAGCCCCCT
 TCATGGGAAACCTGTGCTCCATCCAATGGATGGCTGTGAGGATCCACTCTGTATAGAAT
 CTGGTCAGTGCAGTCCTTGTATCTCAAGAACACTGGGCTGAAAATTAAACCCAAAG
 AACTGTTTTTGTATGATTTGCTGCACTCTTCAATTCCAATAAAGAGTAAGCATCTCAT
 TCCTTGTCTCCCTCTCACTGACCTACCCCTGCCCTTGTGCCCTCTCAAAGAATCAATAAA
 ACCAAAGTGTATAGATTCACTGGCATCTCTCAACTGCTACATCCACTCCAGTAGTATCTC
 ACTTGGCAGGTGAAAAGCTGGAAGCAGTCACGAGGAGTTACAGAAACITAGCCTC
 CTGGAAACCTTGGCATCTCCATAGCTAGAATGCCAGATTGTCCCTGAGATATTGTGGTG
 GGTCTGCTGATGCTTCTGCACTTACTGGATAAGAGTTAGAAATCTCAGGGCAGCT
 TAGCAAAAGTATACTAGAATCTCATGACAGTCAGGTATTGCAAACACTACATTGCAATT
 GAAGAAAGTGTGAAAATTCTCTGACAAATGGAGATTCCCTACAGATAACTAAAGAAC
 AGCTAAGTCACACTCATATGCAAGAATTACCAAGGCTAGGAAAGGGGGGGGGTACT
 GCTTATTCTCATGATAAGGTCTGCTAGAGCAGAACCCCTGGCTGCTACTCACAAGGTC
 AAAGGTGAGCTAAATTGTGACTAGAGTGTGAAATCTTACCTGTGATTAGCTGACTCTA
 GGCAGAGCTTCTTACTCTTACTGTAACCTACCTGGTCTGCTGACACCCCTACTCCCTTGT
 TCTCTGTTTGTGACAGATACTCTATGTAACCTGCCCTGCTGACACCCCTACTCCCTTGT
 TTCTGTATTATATAAGCTGGTGTCCCTTGTGAAAAAATTACATCCAGATACAGCACTC
 CCTTGTGCTGTCCTTGTCTATTTCTGGCCAACCTCATGCCACCTGCCAGAACCCCTA
 GTCTTCTCACAGATTGGGGAGGCCACTGAGCTGGTCCATGGCATCTAACCAACTGTCA
 GCTCACTGTTGGTACTACCTCAAGGTACAAGCTCCATTACTAATGAAACAAAATTAGAT
 AAGTGTGGTCCAGGAAGCAGGGTGTACACCCCTGCTGAATGAAACATTATGAAATGACTG
 AAATAAGTTAACCCATCTCTCTGTTGCTAATATAGCAAATAACCGAGTTCTGAGC
 TGCTGCTGGTGTCTCATGAGGGAGGCCACTGAGCTGGTCTGATCTAGCTTCTGTATGTG
 TGTCCATTGTTCTCATGTTCTGTCCTGGCCCATAGGAAATCTAAGCCATGAAAGGCCATG
 AATCTGGGAATGACTTTCTAAGAAATGCCACGTGAACCTTGGTTCAACGTTTGCCTG
 TAAACAGATATATGGTGCCTGAGTTATAATCATAATAAGCTTGTGAAATAATATAAACCT
 CATTCTCATCTGCTTCCACCGTGGAGCTGAGCTCCCTGGGAGCCACAAAGGCAATTG
 AAAGTAGTGTATTATGACCTGCTTGTGTTACTAATAAAAGCTCATGAAAGGGCAGCC
 ATGTGAAACATGGAACCTGGGGTGCACGTGATCTGTTCTGGGTGCTGACTCATCATA
 TTGCTCCAGAATAATGAGTTATCAACTCTGAGGAAAAGTTGTGTTGTATAGCA
 CGCCCGTGGAGTOCCACCATCTACTCTCTGTAATCTGATATGGTAGAAAAGTTAAATT
 ATGTGATTCTCCAACCTCAAATATTCTAAATCTTCTGCTGCTGCTGCTGCTGCTG
 ACTAAGTCTATTGAGATCTGAGCTGGTATTGCAAGTCATGATGTTCTGATGTTCTG
 TCAGGCTAAGCTGAAAAATGACAAATGAAATGCAATGTCATGTCCTGGGATGTTG
 AGTGTAGGGGTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCCTGGGTT
 CTATCTAGTTGAAATAATGTCCTCTGCTATACAGCAATTAAATGAAATGAAATGCTC
 TTATCTGACCAATCTCAGTGCATCTCATCCGCTCCAAAGTCATGTTCTCATATTCTAAGGAT
 AGGTCTCATTCAACCACATGTTATCATGTTGATGCTGATTTCTGATGCAAATGATT
 ATTCTCCAGAGCACTGAAATTGGGTGAATCATCTTATAACGCCAAAACCTAAATGCTC
 TGCTGCTAAACAGAGAGTTACAAGACCTTTATGTTGATGGCAGCATTAGTCATCCCTA
 TGACAGAAATGTCAGGTGGAGCTCCACTGGGGAGGGCTGGCTCTGGCAGGATTCTC
 TTGGAACATCACACAAAGAAATTCAAATTATGAAATGCACTGATCCATCCAGAATGIG
 ACTTTGACTCTGAAATGAGCTTTAAAGTACGTTGGCTGTCAGACCTTGACTTTGAG
 GTGAAGGAAAGCTGCCAACACTCTCTTATGTAACACAATATGAAAGATCTAATGIG
 GACAGTATGCCAGTCCCAGATCTGTCATGACTGAAAGACACATTGCGATGTTATCACT
 AAGGAGGAGAAGGCAAGCTACAGTGAAGCCCTGTTACTAATAGCTTATGAGAAATT
 AGATAAGAAGGGTTCTAATTAAATTTTTATTAGATATTCTCTCATTTACATTTC
 AAATGCTATCCAAAAGTCCCCTATACCTCCCCCTCACCTGCTCCCCCTACCCACCCACTC
 CCATTCTGGCCCTGGTTACTGTCAGTGGCTCATATGATCCACCAAGCTTACATGCTC
 ACTATCTGGCTATTGCAAGAATGGCTGCGAGCTGATGCGAGTCAGATAACAGACACCTAC
 AGCCAAACAGTGGAAAGGAACCTGGGACTCTATGGAAGAAAAGGAGGAAGGGTTATGG

CCCCGGATGGGGAAAGGAACCTCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
ACCCCTGGGCTCTCAAGAGTCTGAACCACCAACCATAAGAACATTCACTGGCTGTACCCAG
GCCTCTOCACTCATATGTAACAGATATGTGGCTTGGCCTTCATCTGGTCCCTGAACAACTA
GATGGGGGTAGGGGTGGGGATGGGGITATCTCAAAAGCTGTTGCCGTATGTGGGATA
TGTCTTACTGAGCTGCCTAGTCTGGCTCAGTGGAGAGGAAGCACCTAGCTTACGCTT
GTAAAGACTGAACTCTGAGGTGTGGGGAGGGTATACTCAGGGAGGGCCTCACCTGC
TAGGAAGAGAAGAGGAGGGGGAGAGCTTGGGGAGGGGGCAGTGAGCAGGTGGTAAA
ATGAATAAGAaaaaaaaaaaaaATAATTaaaaaaaaaaaaAGAATGGCTGCTGAGCCC
TACTCTAAAACCATTCGATCCCCCCCCCCCCAATCATTCACTGACTACGAATTAAATCATT
GATACTAACATAAGATGTAGGAACATATGTTAATTCTTGTGACTACAGAGTGGTATTG
GAACCTTTTATTGAAGCTTACACAGAGCTGTCTTTCATTCCCTGTACATGCATG
TAGCTTAATGATGTCACTGAAATTAAAACAATGAAGAATAAAGACAACGTATTTT
AAGGATTCTCGTATATATTAAAATCTAAGGTGGTCACCTGGAGAAAATGTCTCAGTT
TTCTATATATGTCTTACTCTATGTTATTAAATTATATGCAATAATTCTAAATCTA
CAACATAGTATGTAACCTATAAGAAAGTAAAACATTCACTGAAATTGTGAAGGTACTTT
CTACOCTCAGAAACACTGGGTTTGAATAATTCTTATTGTTATCAGTGAAATTGAA
AGAATGTAATAACCTACTAAGGCAACATAGAAGGTGAAATTAAAAGAGTAGACAGGA
GAAGTAATAAGGCAAAATAATGAATTATTGCTTAAATAGTCTTAAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAATGACTCTAACCAACAGTCA
TTGAAACAAATGTGTTGAAAATAATAAACATTCTCCTACTTGAAGGATTCTCCTACT
TTGCTTTTATTGCTGATAAGCACCAGACCAAAAGCAGCTT

(SEQ ID NO:18)

72/90

Figure 49

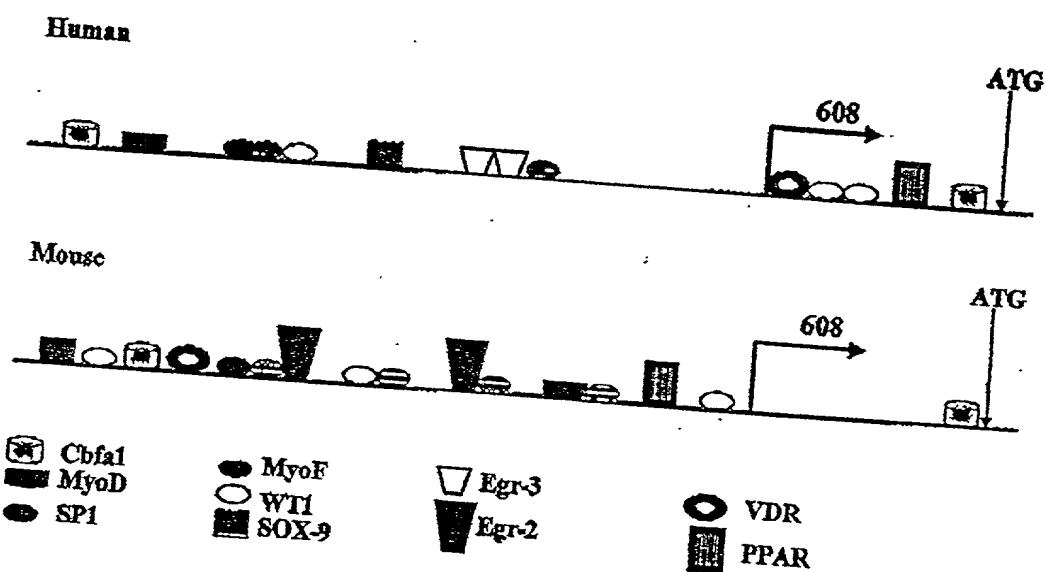


Figure 50

TAAGCCTTGTCCCCTTGGAAAGCAAAGAACGTTCTCAATCAGGTGAAGGCTCT
 CCTCAGAAGATTCTATGTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
 AGAGCTCAGGTATAGCCAACCTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
 CAACTCATGTGGTATTCTATGCGGCTCTCTCTATAACTAACTCTTAAGGTGCA
 TATAGTCTCTCTGTTCCAGCTACCTGCACCATCTTGTTATCTAATAATAGCA
 AGCTCATCTGCTTTAATCATCACGCAGAGAGTATTCAAAAATTCAGTGATGTA
 ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTAAACT
 CATTCTAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGGAACCAACCCAGATGCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
 TTTATCCAATGGAATACTACTCAACTTATTAACAAACGACTTCTATAAAATTCTTAA
 GGCAAATGNATGGTCTGNAGGATCTGAGTGAGGTAACCCAATCACAAAAGAACAC
 TCATGGTATGCACTCACTGATAAGTGGCTATTGCTATGGAGTGATTAAAAGGGA
 AGAAGACACATAGCTTTGTGTGTATAATATTAAGATGAAATTGCCAGTGCTGT
 TTGGCTTATGAGTGAATCTGTTCACTGGATTACCGGAAGAAAATAATAAGTGAAC
 TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
 CCACAGTCCAGGAAACTGAGTCTAAGGATTATTAACCTCCAGTTCCCATGTGC
 AGCTCTGAGACTTGGCAGATCAGACACTTAACCTCACCAAGCTTCTACACAGAGCA
 GTTACTATCCTGCCTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
 GACTTGTAAATAATCTTAAATCCAATTGTGTGAAAGTATGTGACCAAAAGAGCAT
 GGTCTGCTATTAAACCTTGTGTTCTATGGACTCTTAATTGTAGAAATGTCA
 ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCTTCTAGTTG
 AAAGTAATTCACTGTTCTGGCATCCCTTCTAAAGCAATTAAACCCCCAA
 GTAGGCATAATTAAATGCTTACTTCATCAGAATATATCTAATTGACTCTCTAAAGAA
 GACTTGGTATGCATAGGATCTAAATGTAATGTGATTACTGACATAATAAATAGG
 AGAAAATGAGCTAGAATAGGTATAAAATATGTGCTGGCTTCTAATAGGTCTTATAG
 GTTATATAAGAGGGAAAGGAATATTGAAACATCTAGAAGTAAATGATCCTG
 AGTAGCGATCCTGGAAAATACGTACTCTAACACACTGCAATCATCTCTGTGGTT
 TGCTGGAGCTGAGGTCTGGAAAGGCTGACCTGGTTAGAAATAACCTACCGAACATAC
 AGAGCTATGACGTTAGTCTGGAAAGGAGCTTGGAAAGAATGACAAGCTGTAGCTGCC
 CAGAACATACTAGATGCCATATTCCAAGGCAAGTGTCCACATGCGGACATCTAAG
 AATATGGTTGTCTGCACTGCTAAGGACCTTGTGTCGACACACAGGTCTCCAGGG
 TTAGTGCTAACTCTGACTGCTGACTCTTAATTCTCCCTGATCATTAAATGACTAGA
 AATCACTTGGTATTAGCAACTGGATATGGAATTACTTAATTGTACCCAAGCCA
 GGCCACCTCAGCTTGGCAGCTCCATTCTGTGGAGGCCAGTCACGTGGTTTG
 AATCAACTGTACTGTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTCTTCTT
 CACAAGCTTTAGAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGAATGTGTCATGTATGCTCATGGATGCTATTATAATGTGGAAA
 TAAACTGACTTAAAAAAAGCTTATGATACTTGTACAGAGTAAATCTCCATA
 AATATCATCTGCATTATAAATTATTTCATAATCCATCAATTAAAAACCTTAGAAA
 TTTGTTAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
 TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTCAGAACATCTGC
 CTCAAAACATTATCCAAATGCTCATCAAAGGCTCACTCACATGTGCTCAACCAC
 AGGGATTAAACAGTCATTTAGTCACATTCTCAAACGGTGGAACGCTGCTAGAGGA
 ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCTAACCCCTACCA
 CAACAAACAACGACAACAAACCCATAAAAATTATCACGGCAAATGAACTAAGC
 CATATGCAGAAAAAGTATTATATGTTCTCATGTGGGGTGTCTTCTTAATAGTCAA
 ATATGCAGAATATAGACAAAGATGGTTATGCAAGTGGGGATGGCGAAGGATACTT
 GTAGATTAGAGGACACAAAGCAACAACACTACAGAGTGAAGTAATCCAGAGACTAAT
 GTATAATATGAGGACTGTATTAAATAATTCTATTAAAGATACACAGCAAACGAGTGT
 ATCTTACTAACACACACACTACATAGAGAGATAAAAGTGTAGATACGTTGTTT
 ATCTTCATGTAGCTGATAATTCAATTGTACACCTCAAACATAGATAACCAACAAA
 GAGGAAGAGGATAGGTGCCTCTCCAGGGCGGAAGAGTACATTGAAAGTCAGACA
 CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
 GGATCTGCAACCCCTATAGGTGGAACAACATTATGAGCTAACCACTGACCCGGAGCT
 CTTGACTCTAGCTGCATATATATCAAAGATGGCTAACGCCATCACTGGAAAGA
 GAGGCCATTGGACTTGCAAACCTTATATGCCCAAGTACAGGGGAATACCAAGGCCA
 AAAAGGGGGAGTGGTGGCAGGGAGTGGGGTGGATATGGGGACTTTT
 GGTATAGCATTGAAATGTAAATGAGTAAATACCTAATAAAAATGGAAAAAAA
 AAAAAAAAGGAAGGTCAAGCACACCTCACTTCACTGCTATCTCAACTTG
 CAAACAGAAGGGAGTCACAAACCCAGGACAAACCCACAGTGATTGAAGCGTCTTG
 AATGTTATTGCTGTTGTTACCAACATCATTAGCATATATTCAATTGTGAAAACCTA
 CGGGGTCTATGACATGTTTTTATTCAAGTATATCACATGCTGTCAGCATATTGGC
 ACCACTACCAGCCCCAGCCCCCTTGCCCCGCCAACACACACACACACACACAC
 AC
 TCATCTGCTCACTCACCCACCCAAAGCTTAATCCTTCTCCCTGCAATAGTACCTC
 TCCTATTGTTATGCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTACGA
 AAGAAAGGAATTGTAGCCCTGAGCCAGCCGGGCCACAGAGCCTGCCACAGAC
 AGGAAAAGCCCAGGGCTTACCAAGCAGAGGAGGAGCAAACACTCGCAGGCGAGCCTGG
 GTTGGCGCTGGTGGTCCCAGGGCTGATGGCCGCCATTCCCAGAACGCCAGGCTATA
 GCTGCGTCACCTGCCCTCCTCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
 CCCAAAGGCGGTGCAATTCCATTGGCCAAGGCAGAGGTGAGCCTGCTAATCCC
 CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTCAAGGG
 CGAGACAGAGGGTGGATGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGGG
 AGTGGGACAAAGCAGCAGCAAGGGAACATCGCTTCAGTACTGAAGCCAGGCA
 AAAGGAGCGGGAAAGGATTATATGTAGCCTGGACGCTTCATAAACACTGATGACG
 TGTTGTGCAAAGCAAGCAATTGAGGAGAAACGCCCTGGACGTCGGAAAGAAGGA

GTGATCGATTAGTACTTGTAAAGTTAGGTGAGTTGAGAACTAACA
 TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAATCAGATAA
 AGGAAAGCTTGGTTAGTTGGAAATGTATGATACCATTAAAATAACAGAAGCGCCT
 CCAGTTCTCTGAAGAGTCAGTCCCCAGCTAGTGAAGACTAACGCCTACTAACGCCTT
 TGCTCCCGTTGGAAGCAAAGAACGTTCTCAATCAGGTGAAGGCTCTCCTCAGAAG
 ATTCCTGTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAG
 GTATTGCCAACTCTTGTAAATACAGTTGAGGCTTAAGTGTACGGGAACACTATG
 TGGTATTCACTTACGGCTCTCTCTTATAACTAACCTTAAGGTGCATATAGTCTC
 TTCTGTTCCCAGCTACCTGTACCATCTTGTATTCTAATAATAGCAAGCTCATCT
 GCTTTAAATCATCACGCAGAGAGTATTCAAAAATATTCAAGTGTAAACAGTGACA
 GTGTAGGCATAGAAGTAATCATTAGTAAATCTAATTGGTTAAACTCATTCAA
 CAGCTCCAGGTTGGGAGGGATCACTGAGCCTCGCCACGTGCAGGGTAAAGATATT
 TCTAACAAAGAGAAGCAGAATTCTCCTGGCCATGCTCCCCATCACTGTGTCAGTAA
 GCAGAGGGGTGTTCCAAGCAGAGAAAGAGCAGACAGTGTATGCCTGCAAAGTCA
 GAGACTCAGCCCTCCCAGCTGGTCAGTTACTGTCCTCCGGTCATTAGTGGCTCTG
 AAAAGGCCATGTGCTTATTGCAAGGACTTGCAAGACATGCTAGAAAGAAATT
 GACCTTTCTAGTGGTTATTACAGCTGTAAAAGTATTGGAAGGTTAAGCCA
 AATAAAATAAAACACATATTAAATAATACAATGTTACAAAATTGATCATATAAAGA
 AGTACATTCAAAATGCAATGTAAAAATATATAATTTTATCTATTACTGGTGC
 AAAGTTTCTAAATTGCACATGTACTATTITATTTATAAAAATATTAAAGTCA
 TATATAAAAGTGTAAAAGGCTTGGTCAAACAAGAGAGTAAATTACAAACTTAA
 ATTGTCCCCATAACATTATTATGATCTCTAAATGACAGGGATCTGCTTTCAATTGGG
 AATGAGAAGCTATGAAGATATGTTACAATAATAAGCCCATTAGTGTAAAGTCCA
 ATGGGAAGCTAGCACACACTGGTTATAAAGAGAACAGTGTCTGAGTCTATGCAA
 GTTTACACTCTAGGAATAAGAGTTCCCTTCTCCAGATTCACTAGCATTGTTG
 CATCATTATCTTGTGATGAGCATTATAAGGAAATAAGATAGGATCTCAAAG
 GAATGTCAATTGGATGCCCTGAACAATCTTCAAGGTCTTCTTCAGTTCACTAGTC
 TATTCAATTATTGATAATTGGGGGATGGTGTAAATTGTCAGTTCTATGGAA
 TTCCAAAAAAACAAAAACAAACCAACCAACCAACCAACCTCTGAAACTAGAAACTACC
 AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAACATTATTGCTACATTG
 TTCATTATTCAACAGCCAAGAACGAAATGTGAAACCAACTTACGTAGCCGTCAAATATG
 AACGGATAAAAGAAAATGTGAAATGTGTACAACAGAGTCCCAGTGGCCATAAAAG
 AGTAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGCTAAT
 CAAAACAAGACAGACTCAAAAGGAAACACCGGTAGCTTCTCTGACAAACAGAAG
 CTAGATTACACTGTACGTGCGCATGTGTGTTAGAATTATTAGTTACACTA
 TTCTAAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAACAAAGCTAGCTG
 GGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
 ATAGGTGCCAGGATGAAATGCATTAATTGTATGCTAACTAAACCACAGACAGGAG
 GCACACGTTCAAACCAAGGGTGAATCCCAGCACAGAGAACGGGAAGTAGACACAA
 AGTTGCCACTAACCAAGAACCCATTGCAAGTTGCTGCCTGCTGGGAAGGGCGTT

CCAGTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
 ATGATGGATGGTTTTGTGCTATTTTCACTTTTCTTACTGTTTGTGTTGTGGT
 GGTGTTGTGGTGGCTGTGGTTTCATTTGTTCTTTGAGAGAGAGAAGGAAC
 ATGAAATTGGTGGTAGGAAGCTGGAAACGATCTGGAAGAAGTGGGGAAAGAG
 AAAAATTGTATGGAGCATATTAAACAAACAAACAAACAAACAAAGGTCATT
 GCCACAAAAAGGTGTGAATTAAATTAAACCAGTACGACTCTTAAAGAAAATATTCCC
 AATTATTCCCAGAGTTGCTATGTATGCTGTGCCAGGACTTGTGTTGAACGGCCCTA
 TAACTCTGGTGTGGTCTTTCAGGATGCAGAAGAGAGGCAGGGAAAGTCAGCTGCT
 TGCTGATCTCCCTCACTGCCATCTGCCGGTGGTCACCCCTGGGAGCAGGGTCTGTC
 CTCGCCGATGTGCCCTGCTATGTGCCACAGAGGTGCAGTACATTGGGACCTGA
 CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTAGGGTGTG
 GACCTTGCCTGATCTCCTCTCAGAGAGGGACCACTGATTTCCTGGTACTTGGCC
 CCAAACACCTGTGATTACTTTAATAGTTCTAAATGGGTCATACAAACCTT
 ATATTGTGGAGACAATGAACATTATCCAATAGTCTTTACTAGAACTTGAAGCC
 CCTCTAGTTGTTGGGAGCCTCATAATTATGGGCAGCTTATTCTGAATGAATT
 AAATGAAAAAGATAACAGTTCTGTTAACATCAATTATGATAACCAAGGAAGAGGAAT
 TGTCTTGAATATTAAAAAGCATTCTTGTCAATTATAAAATACCCATTACAAA
 ATGGCTTACTTAAATACITGCCCTACTAAATCTGACAAATTATGGTATTTGAA
 GGTTATGAAAATTGTTATGTGTATAATGCACAAGAAATGGGATATGCCATCAC
 CTATGTGCCATTAGTGAGCATGTACAGTATGCCAACACTATTGTCACGTTGGAG
 GAAGTAATGGGGTGGGGAGCAACAAGGGTATAACCGTATACCCAGTGCCTGG
 AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGCCTT
 GGGCTGAGCACTTGCAATGAGCATTGCTCATTGTGCTGGCAGGTTATGATAAC
 TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTCAACTGAGAACATGCCTGA
 AGAAGATCAGATTAGGCAGGCCGTGGGCATTTCITAATTAGTGATTGATGGG
 GCAGGGCCCAGTCCATTGTCGTGGTACCATTTCTCAGGCACTATTAAAAAAAAAAA
 AACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
 TGCATCAGCTCCTGCCCTCAGGTTCTACCCATTGAGTTCTGCTCTAGCTCC
 CAGTGATGAACAATGATGTGGAAGTATAAGCCAATAAACTTCTTCTCCCAACTT
 GCTGTTGGTCATGATGTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
 TTATAACACCTTGGACTAAATTCTGTTATCTATAGTGAGGAAAATGGAGCATAGA
 AAGTCTCCAGACTACACCAGAGTGAAATCTGGCCTGAGCTTGAAGATCACACCCAC
 GTGCACTCCACTGCCGGGCTTCTTAACCGGAACACAGTTGAAAGGAATTCT
 GTTGTTCCATTGACATGTGACTTAATTGACGATTCTGAAGCTGAAAGATG
 ATTGCAACCTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTCTGGCCT
 GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTCAAGAGTCAGTGACA
 AGACCTCTCGGGCTGCACTGCCCTGAGGTGAGATAGGTAGAGGGTGTGGAGGC
 TGAGAAGAGAGGTGCAACTGTGGGTTATACCCAAAAGCTGCTGATTCCGTGGAG
 ACATTCTATAAGCATTCTATAAAACTAGAGGCAGATATCAAGGAAGGATTCAATTGT
 AATGCAATTGAGAAAATTGAATATTAAGAAAATGCTGGGAAAATGCTTAC

ACAATTGCGAGGACCTAATTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
GCGGCAGACACCTGCAATTCCCTGTCCTGGAAAGCACCTGTCAGAATCCAGAGACT
CATTGGCCAACACTCTATTCAATCAATGAAGTCCATATTCACTGACAAAACITGAC
TCAGAAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
GCATGAATAAGGGATCCCAGAGAGAAGGGAAAGAAAAAGGAAGGAAGGAAGGAAG
GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGAGGGAGGGAGGAAGGAAG
AGGGAAAGGAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG
GAAAGGGGGAGAAAGAAGAGAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG
ACACCTTAAATTTATCCATAAAAGGTCAATTCCACCTGTTGTCTGGAAGTAGAGT
GGGATCCCTATATAAGGGCAGTCCTAACATAGTAGCATTTATAAACCAATTACAA
ATTTGAGTTTCTACTTTATCCTCTACCATCTCAAACACTGAAACTACAATTATT
CCCACAAATGAAGAAAATGCTGTAAGAGTTTCACACACCGAAGTGGAAACTTAA
GGATTAGACAAGTCTAACATGAGAATGGGAGAACAAAAAGAGACTGCACAGGG
AGCCCTTCTGTTATAATCTTGCACACTGAGAAGCTAATTGACGCTGCATGACTA
CTCAACTCTTAAGCAAACAATGCTGTTCATGAAAAGCACAATAAAAGTACATAT
GTCCCATAATATTCAAAATTGCACTGCAGCACATAATAGCAATCAAAGCAATAA
CACCCACTGTCACAGAGACTTAAACATGAAACTGGAACATGTCTAGTGTGTTGA
CTTAGGGTACATAGTATGCTGTCGTATGTACCAATGTTGATTAGGTCACTCAGA
CAGCATTGGAACATGTATCTCAGGAGGAATCATTGATGTCATGCTGCAATTG
TCCACCTATGTTATTCTCTAGCCAGGTTCTGATGGAGAACATTGGGTTG
AGGTTTACTCCCAGGTAAACATTAGGGAAAAGCTGTCTATGTTCTCAGTTGGCTT
TATTATGAGGGATGTTGTTATTCCAGAAAATTCTCTTTGAAGAGATTACAATT
GGTCAAACAGAAAATATGAAAAAGTTATTGTTTATTAGTATTGATGTTCTT
TCTTTTAAATGGTATGCTTAGAAACTAATTAAAGATTAGATTAGATTAGA
AAATAATCAGAGAGGGATTGATGAATGCTAACAGCATGAAAAATTCAAATT
TTGCTCTAATTCAATTGAGAACATTAAATTCTATTACTATAAAAGACAGCACGCCAG
ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGAGAAC
GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTGAAAACATTGGCTACATG
ATGGAAGCCAGGCACAAAAAGCCACATATTGATGGTTATGTTATATGAAATGTT
AAAATACATGGATTCTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
TTAAAAAAACTATTGATGTGGAATGATCTTAATTGGGAAAAGACAATT
TCCTAAGACGAAATAGTGTAGGAGATAGTTATATCCCTGTTGATATTGAAATAA
ACCAGCATGCTGCTCTGAGAAGGGCTAATGAAGGGCAGGAGGAAGTGAATG
AGATGGTAGAAAGGAAAGTCATATACCATGGCTCTCGTGGTGGAAATCTAGAT
ATGTTAATATATTGACATAAAGGAAGGAATTGTTAGGGAGGATCAAACCAACA
GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTCTGCAATGTGTT
GGAGACACCATAATAAAACTCCTTTGTTGCTAACTAAAACCACTAAATTCTAA
AAACAAAACATTTCGACAAGAATTATTATTCAATAAAAGATGTTAAATGGG
GGAAGTTGAAGTCATTGATAGTCTCATAAACTTAAATGTATTAAACTGCTT
CGTTTTTATTATTAAATTACTCTTGTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTCACCACACAAATGTAGGCAGAAGAGTGTAACTCCACTTAG
TGAGGCAATCTGGAGAGGGAAAGGAAGCGGATCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTGTATTCTAATATTAAGAAAATAATAAAACCCATTCTGTG
CACTTATCACCCAGGCTAACAGTTATCTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTAAAAGTAACCATGACATCTCTGTTCCT
TTAGCGATAAGTTCTAATATTATGGCTGCTGTATGTTCTAATTCTCTAATATT
GTCACATTAGTGGCAACTACTTGTGAATTGAGTGGAGTTAAGGTCCCAG
GATTAATCTCAACATATTCTATATTATAAACTTTCTCTTGTGAAAGTTCCCTT
GAGAAAACAAATATGCCATATCTTCTTACAGGTCTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGAACTTTGTATGGACTCAGGAGCTGACCCGGTTGCA
CCTGGATCACAAACATTGAGTTATCAACCCCGAGGCCTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTGAGCTATCTCCAGATATTAAAACCTCCTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCAAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACTATCTTGTGTTGTTCTTTTT
ATARKACGTATTTCTCAATTCTATTAGAATGATATCCAAAAGTCCCCCATAACC
TCCCCCCCACCTCCCTACCTACCCATTCCCATTGGCCCTGGCATTCCCTGTACT
GGGCATATAAAGTTGCGTGTCCAATGGACCTCTTCCAGTGATGGCCAAGTAG
GCCATCTTGATACATATGCAGCTAGAGTCAGAGCTCTGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTGAA (SEQ ID NO:20)

Figure 51

MPKRAHWGALSVVLILLWGHPRVALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
 RINLGFSNIQALSETSFAGLTKELELMHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
 QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSRLLHLEGNLLHQLHPSTFSTFTFLDYFR
 LSTIRHLYLAENMVRTLPASMLRNMPLENLYLQGNPWTCDCEMRWFLEWDAKSRGIL
 KCKKDKAYPEGGQLCAMCSPKKLYKHEHKLKDMTCLKPSIESPLRQNRSRSIEEEQEQ
 EEDGGSQQLILEKFQLPQWSISLNMTDEHGNMVNLVCDIKKPMVDVYKJHLNQTDPPDIDIN
 ATVALDFECPMTRENEYEKLWKIAYYSEPVKLHREMLSKDPRVSYQYRQDADEEAL
 YYTGVRQAQILAEPPEWVMQPSIDIQLNRRQSTAKKVLLSYYTQYSQTISTKDTRQARGRS
 WVMIEPSGAVQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPSKFSIL
 SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVRVLVQSPSTQPAEKDTVTIGKNPGES
 VTLPCNALAIAPEAHLSWILPNRRJINDLANTSHVYMLPNGTLSIPKVQVSDSGYYRCAV
 NQQGADHFTVGITVTKGSGLPSKRGRRPGAKALSRVREDIVEDEGGSGMGDEENTSR
 RLLHPKDQEVLKTKDDAINGDKAKGRRKLKLWKHSEKEPETNVAEGRRVFESRR
 INMANKQINPERWADILAKVRGKNLPKGTEVPLIKTTSSPPSLSLEVTPPFPAVSPPSASP
 VQTVTSAESESSADVPLLGEEEHVLTGTTTSSASMGLEHJHNGVILVEPEVTSTPLEEVVDDL
 SEKTEEITSTEGDLKGTAAPTLISEPYEPSPTLHTLDTVYEKPTHEETATEGWSAADVGSS
 PEPTSSEYEPLDAVSLAESEPMQYFDPDLETKSQPDDEDKMKEDTFAHLLPTPTIWNDS
 STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKGMKEMSQTQGGNM
 LEGDPHTHSRSSESEGQESKSITLPDSTLGIMSSMSPVKKPAETTVGTLKDTTVTTTPR
 QKVAPSSTMSTHPSSRRPNGRRRLRPNKFRHRHKQTPTTFAFSETFSTQPTQAPDIKISS
 QVESSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGTPRRKHGKRPNHRYTPSTVSSRA
 SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDYMTTRKIYSSYPKVQETLP
 VTYKPTSDGKEIKDDVATNVDKHKSILVTGESITNAIPTSRSLVSTMGEFKEESSPVGF
 GTPTWNPSRTAQPGRLQTDIPVTTSGENLTDPPPLLKELEDVDFTEFLSSLTVSTPFHQEE
 AGSSTTLSSIKVEVASSQAETTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSPS
 TILMSLGTTTITKPALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEGTQHMSGPNE
 LSTPSSDRDAFNLSKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPLPTA
 TVRLPEMSTQSASRYFVTSQSPRHWTKPEITTPSGALPENKQFTTPRLSSTIPLPLHM
 SKPSIPSFKTDRRTDQFNGYSKVFGNNSIPEARNPVGKPPSPRIPHYSNGRLPFTNKTLSF
 PQLGVTRRPQIPTSPAPVMRERKVPGSYNRISHSTFHLDGFPPAPPLLHTPQTGSPSTN
 LQNIPMVSSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
 TAETDTVPCATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
 QYMCTASNLHGLDRMVVLLSFTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
 SWIFPDRRVWQTVPESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
 AALPPVIHQEKLENISLPPGLSIHCTAKAAPLPSVRWVLGDTQIRPSQFLHGNLFVFP
 NGTLYIRNLAPKDSGRYECVAANLVSARRTVQLNVQRAAANARITGTPRRTDVRYG
 GTLKLDCSASGDPWPRLWRLPSKRMIDALFSFDSRIKVFANGTLVVKSVTDK

DAGDYLCAVNKGDDYVVLKVDVVMKPAKIEHKEENDHKVFYGGDLKVDCVATGL
PNPEISWLPDGSLVNSFMQSDDSGGRTKRYVVFNNGTLYFNEVMREEGDYTCFAEN
QVGKDEMVRVVKVVTAPATIRNKTYLAQVVPYGDVVTVACEAKGEPMVKVTWLSPN
KVIPTSSEKYQIYQDGTLIQLKAQRSDSGNYTCLVRNSAGEDRKTWVHVNQPPKING
NPNPITTREIAAGGSRKLIDCKAEGIPTPRVLWAFPEGVVLPAFYGNRITVHGNGSLDI
RSLRKSDSVQLVCMARNEGGEARLIVQLTVLEPMEKPIFHDPISEKITAMAGHTISLNC
AAGTPTPSLVWVLPNGTDLQSGQQLQRFYHKADGMLHISGLSSVDAGAYRCVARNA
GHTERLVSLKVGLKPEANKQYHNLVSIINGETLKLPCPTPGAGQGRFSWTLNGMHLEG
PQTLGRVSLLDNGTLTVREASVFDRGTYVCRMETEYGPSVTSIPVIVIAYPPRITSEPTPVI
YTRPGNTVKLNCMAMGIPKADITWELPDKSHLKAGVQARLYGNRFLHPQGSLTIQHAT
QRDAGFYKCMAKNILGSDSKTTYIHVF (SEQ ID NO: 21)

Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT
 TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT
 ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCC
 GCTGGCATTG CTAGACACGT GGAAAGAAC AATTGGGGT TTAATAGCAT
 ACAGGCCCTG TCAGAAACCT CATTGCAGG ACTGACCAAG TTGGAGCTAC
 TTATGATTCA CGGCAATGAG ATCCAAGCA TCCCCGATGG AGCTTAAAGA
 GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT
 GATCACAGGA CAGACCCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA
 TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTT CAACGGCTTA
 ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA
 CCCCAGCACC TTCTCCACGT TCACATTGGT GGATTATTC AGACTCTCCA
 CCATAAGGCA CCTCTACTTA GCAGAGAAC TGGTTAGAAC TCTTCCTGCC
 AGCATGCTTC GGAACATGCC GCTCTGGAG AATCTTACT TGCAGGGAAA
 TCCGTGGACC TGGGATTGTG AGATGAGATG GTTTTGGAA TGGGATGCAA
 AATCCAGAGG AATTCTGAAG TGTAAAAAGG ACAAAAGCTT TGAAGGCGGT
 CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT
 ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCCTC
 TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG
 GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAAGTGC CCCAGTGGAG
 CATCTCTTGT AATATGACCG ACGAGCACCG GAACATGGTG AACTGGTCT
 GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG
 GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCTGG ACTTTGAGTG
 TCCAATGACC CGAGAAAAT ATGAAAAGCT ATGAAATTG ATAGCATACT
 ACAGTGAAGT TCCCCTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC
 CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTAA
 CTACACAGGT GTGAGAGCCC AGATTCTGC AGAACCCAGAA TGGGTATGC
 AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG
 GTGCTACTTT CCTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA
 TACAAGGCAG GCTCGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG
 CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC
 TGCAACGTGA AAGCTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA
 TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA
 TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA
 GGCTTGTACC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT
 ATATAAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGGCCAGCC GAGAAAGACA
 CAGTGACAAT TGGCAAGAAC CCAGGGAGT CGGTGACATT GCCTGCAAT
 GCTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG
 GATAATTAAT GATTGGCTA ACACATCACA TGTATACATG TTGCCAAATG
 GAACTCTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTACGG
TGGGAATCAC AGTGACCAAG AAAGGGCTG GCTTGCCATC CAAAAGAGGC
AGACGCCAG GTGCAAAGGC TCTTCCAGA GTCAGAGAAG ACATCGTGG
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
GAAGCATTG GAAAAAGAAC CAGAGACCAA TGTGCAGAA GGTGCGAGAG
TGTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
GAGCGCTGGG CTGATATTTC AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
GGGCACAGAA GTACCCCCGT TGATTAAAAC CACAAGTCCT CCATCCTTGA
GCCTAGAAGT CACACCACCT TTTCCTGCTG TTTCTCCCCC CTCAGCATCT
CCTGTGCAGA CAGTAACCAAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
ACTTGGTGAAGA GAAGAGCACG TTTGGGTAC CATTCCCTCA GCCAGCATGG
GGCTAGAACAA CAACCACAAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA
AGCACACCTC TGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
TATCTGAGCC TTATGAACCA TCTCTACTC TGCACACATT AGACACAGTC
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
AGATGTTGGA TCGTCACCAAG AGCCCACATC CAGTGAGTAT GAGCCTCCAT
TGGATGCTGT CTCCCTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
GATTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
CTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTTAAT GACTCCAGTA
CATCACAGTT ATTGAGGAT TCTACTATAG GGGAACCAAGG TGTCCCAGGC
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTG TGAAAAGTAG
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCTGA
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCAG
AAACCACAGT TGTACCCCTC CTAGACAAAG ACACCCACAAG AGTAACAACA
ACACCAAGGC AAAAAGTTGC TCCGTACATCC ACCATGAGCA CTCACCCCTC
TCGAAGGAGA CCCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
ACCGGCACAA GCAAACCCCA CCCACAACCTT TTGCCCCATC AGAGACTTTT
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTCAA GTCAAGTGG
GAGTTCTCTG GTTCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
CGGAGAAAAC ACGGGAAAGAG GCAAACAAA CATCGATATA CCCCTCTAC
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
AACATAGAAA CATTGTTACT CCCAGTTCAAG AACTATACT TTTGCCTAGA
ACTGTTCTC TGAAAAGTGA GGGCCCTAT GATTCCCTAG ATTACATGAC
AACCACCAAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
TTGACAAACA TAAAAGT GAC ATTTAGTCA CTGGTGAATC AATTACTAAT
GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTAAAGGA
AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
GGACGGGCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
TTTCACTTCC GAGTTTTGT CCTCTTGAC AGTCTCCACA CCATTTCACC
AGGAAGAAGC TGGTTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
GCTTCAAGTC AGGCAGAAC CACCACCTT GATCAAGATC ATCTTGAAAC
CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA
CTGCTGCCG GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
ATGTCTTGG GACAAACCCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG
AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTC TTGAATTATG
TGGGAATCC AGAAACAGAA GCAACCCAG TCAACAATGA AGGAACACAG
CATATGTCAG GCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
ATTAACTTG TCTACAAAGC TGGATTGGA AAAGCAAGTA TTTGGTAGTA
GGAGTCTACC ACGTGGCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
GCTTCTCATC AACTAACCAAG AGTCCCTGCC AAACCCATCC TACCAACAGC
AACAGTGAGG CTACCTGAAA TGTCACACAA AAGCGCTTCC AGATACTTIG
TAACCTCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
TATCCTTCTG GGGCTTGCAGAGAACAAA CAGTTACAA CTCCAAGATT
ATCAAGTACA ACAATTCCCTC TCCCATTGCA CATGTCCAAA CCCAGCATT
CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTACTCCAAA
GTGTTGGAA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC
TCCCAGTCCA AGAATTCCCTC ATTATTCCAA TGGAAGACTC CCTTCTTITA
CCAACAAAGAC TCTTCTTCC CACAGTTGG GAGTCACCCG GAGACCCAG
ATACCCACTT CTCCTGCCAGTAATGAGA GAGAGAAAAG TTATTCCAGG
TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC
CTCCGGCACC TCCGTTGTG CACACTCCGC AGACCACGGG ATCACCCCTCA
ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
CTTATAACA TCTTCTGTCC AGTCTCAGG AAGCTTCCAC CAGAGCAGCT
CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTTTGGG
GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
AGGATACAAC GGTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
GGTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
ACGGCCTGGA CAGGATGGTG GTCTTGCTT CGGTACCGGT GCAGCAACCT
CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCC CAAATTCCCT
GGATCTTCCC TGACAGGAGG GTGTGGCAAAC CTGTGTCCCC

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCCT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GGCCTCTATA AGTGCCTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCC
CGTTATCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCC CCGGGGCTCA
GCATTACAT TCACTGCCT GCCAAGGCTG CGCCCTGCC CAGCGTGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCG
CCAAGGACAG CGGGCGCTAT GAGTGCCTGG CCGCCAACCT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCG
CATCACGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACCGC CTCTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TCGTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA
GACGGGAGTC TGGTGAACCTC CTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTTAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTGTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCACCTCTCC TTATTCAAGAA AGCCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATTC ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGT ATGGGCTTTT
CCCGAGGGTG TGGTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGCTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGCCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTCG GTCACCAGCA TCCCCGTGAT TGTGATGCC
TATCCTCCCC GGATCACCAAG CGAGCCCACC CCGGTACATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCCAGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACCT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTGTA
AGGGAAAGCCA GTTGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTCACTCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTACA TGATAGACTT TGTTCCAGAT TGACAAGTCA
TCTTCAGTT ATTCCTCTG TCACCTCAAA ACTCCAGCTT GCCCAATAAG
GATTAGAAC CAGAGTGAAT GATATATATA TATATATTAA ATTCAAGAGT
TACATACATA CAGCTACCAT TTTATATGAA AAAAGAAAAA CATTCTTCC
TGGAACTCAC TTTTATATA ATGTTTATA TATATATTAA TTCCTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATAAAATTA
ATAAATTATT GGTCTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTAA AAAAATTTC TCTCCAACCT CCTTCAAATT CAGTCACCAC
TGTTATATTA CCTTCTCCAG GAACCTCCA GTGGGGAAAGG CTGCGATATT
AGATTTCTT GTATGCAAAG TTTTGTGA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAAACTTC TCTGCAGTAT CTGGCTTGTGTC
CATCTGGTCT AAGGTGGCTG CTTCTCCCC AGCCATGAGT CAGTTGTGC
CCATGAATAA TACACGACCT GTTATTCCA TGACTGCTT ACTGTATTAA
TAAGGTCAAT ATACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAAA AAAAAA

Figure 53

Human OCP: nucleotide sequence of ORF

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTGGTCTCCTTGCTGTGATCTGC
 CTGGTCGCCACCCCTGGGGCAAGGCCTGCTCGCCGCTGCTGTTATATGCCT
 ACGGAGGTACACTGCACATTCCGGTACCTGACTTCCATCCAGACAGCATCCGCC
 AATGTGGAACGCATCAATTAGGATAACAACAGCTGGTAGATTGATGGAAACAGAT
 TTTTCTGGCCTGACCAAACGGAGTTACTCATGCTTCACAGCAATGGCATTACACA
 ATCCCTGACAAGACCTCTCAGATTGCAGGCCTGCAAGGCTTAAAGATGAGCTAT
 AATAAAGTCCGAAAACCTCAGAAAGATACTTTTATGGCCTCAGGAGCTGACACGA
 TTGCACATGGACCACAACAATTGAGTTATAAACCCAGAGGTTTTATGGGCTC
 AACTTTCTCCGCCTGGTGCACTTGGAAAGGAAATCAGCTCACTAAGCTCCACCCAGAT
 ACATTGTCTCTTGAGCTACCTCCAGATATTAAAATCTCTTCATTAAGTTCCTA
 TACTGTCTGATAACTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTATATGCCT
 GACCTAGACAGCCTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTAAAG
 TGGTTGTCTGACTGGATACAGGAGAACAGCAGATGTAATAAAATGCAAAAAGATAGA
 AGTCCCTCTAGTGCTCAGCAGTGTCCACTTGATGAAACCTAGGACTTCTAAAGGC
 AAGCCGTTAGCTATGGTCTCAGCTGCAGCTTCCAGTGTGCCAAGCCAACCATTGAC
 TCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTCATC
 TCTCCCCAAGGTTCATGGCACCCCTTGGCTCCCTCACTTGAATATGACAGATCAG
 TCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCC
 ATTGCATTCACTGAAGAAAATGACTACATCGTGTAAATACTCATTTCACATT
 TTGGTGTGCAACATAGATTACGGTCACATTCAAGCCAGTGTGGCAAATTGGCTTG
 TACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTGCTTAGTGAACACCCGCAG
 CTCTATTACAAATATAAACAGGTGGCTCTAACGCTGAAGACATTTACCAACATA
 GAGGCAGATCTCAGAGCAGATCCCTCTGGTTAATGCAAGACCAAATTCCCTGCAG
 CTGAACAGAACTGCCACCACATTCACTGACATTACAGATCCAGTACTCCAGTGATGCT
 CAAATCACTTACCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATT
 TCAAGGGATAACAATACTAAGCTGGAACATACTGTCTGGTAGGTGGAACCGTTGGC
 CTGAACTGCCAGGCCAAGGAGACCCCACCCACACGTGGATTGGCTTAGCTGAT
 GGAAGTAAAGTGGAGAGGCCCTATGTCAGTGAGGATGGACGGATCTAACAGACAAA
 AGTGGAAAATTGGAACCTCAGATGGCTGATAGTTGACACAGCGTATATCACTGT
 ATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAA
 CCTTGGTCGAAGCCTATCAGGAAATGGGATTCATCACACAGTTTCATTGGTGA
 ACACTTGATCTCCATGCCATTCTACTGGTATCCCAGATGCCCTATTAGCTGGTT
 ATTCCAGGAAACAATGTGCTCTACAGTCAGTCATCAAGAGACAAGAAAGTTCTAAACAAAT
 GGCACATTAAGAATATTACAGGTCAACCGAAAGACCAAGGTTATTATCGCTGTG
 GCAGCCAACCCATCAGGGTTGATTTTGATTTCCAAGTTCACTGCAAGATGAAA
 GGACAAAGGCCCTGGAGCATGGAGAAACAGAGGGATCTGGACTTGATGAGTCC
 AATCCTATTGCTCATCTTAAGGAGCCACCAGGTGCACAACCTCCGTACATCTGCTCTG
 ATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACAAGTAAGAGGCACAACATCGG
 GAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTAGGGAGAATAGG
 AGGCATTCCCTCCCTGCTAGGAGAATTGACCCACAACATTGGCGGGACTGTTG
 GAGAAAGCTAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGGCCA
 CCCCCAGTGGTCACCCAACTCCAAACATACCTGGTGAAGAAGACGATTCCCTCAGGC
 ATGCTCGCTCATGAGGAATTATGGTCCCGGCCACTAAAGCTTGAACCTTCCA
 GCAAGGACAGTGACTGCTGACTCCAGAACAAATCTGATAGTCCTATGACAAACATA
 AATTATGGCACAGAATTCTCCTGTTGTGAATTCAAAATACTACCACCTGAAGAA
 CCCACAGATTCAAACGTCTACTGCTATTAAACTACAGCCATGTCAAAGAATATA

AACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCACTCCACTGTC
 TTTCCACTGCTACTTGGAGCAACTGAATTTCAGGACTCTGACCAGATGGGAAGAGGA
 AGAGAGCATTCCAAAGTAGACCCCCAATAACAGTAAGGACTATGATCAAAGATGTC
 AATGTCAAAATGCTTAGTAGCACCACCAACAAACTATTATTAGAGTCAGTAATACC
 ACAAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAAACCCAGGCACAATCACTTC
 TATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTCAGATCCACACACA
 GCTGCTCATTCTCAGTTCCGATCCCTAGAAATAGTACAGTTAACATCCGCTGTT
 AGACGCTTGGGAGGGAGAGGAAATTGGCGGAAGGGGGGGATTATCAGCCCATA
 AGAACTCCAGTTCTGCGACGGCATAGATAACAGCATTTCAGGTCAACAAACCAAGAGGT
 TCTTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAATGTGACATGTCTG
 TCCGTCTTCCAGGGAGAGGCTCACCACTGCCACAGCAGCATTGCTTTCAGT
 GCTGCTCCATCACCTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCT
 ACAACTCTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAA
 ACAACACCCACAATAAAATATTCAAGGACTGAAATTCCAAGTGAACCTCAACTGGT
 GCAGTCATGACATATGCTCAACATCCATACCCATGAAAAAAACTCACAAAGTAAAC
 GCCAGTTACCCACGTGTCTAGCACCAATGAAGCTAAAGAGATTCACTGATTACA
 TCGTCACTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACAAGG
 TTTCAAGAAGGAAATTCCCTGGCAACAGAACCTTGTAAATAACCATAACCCAAAA
 GGCAGATTAAGGAATCAACATAAAGTTAGTTACAAAAAAAGCACAGCTGTGATGCTT
 CCTAAAACATCTCCTGCTTACACAGAGACAAAGTCCCTTCCATTCAACCA
 CTTCAACAAAGTGTGATGCAAATTCCATCTAACCTTGACTACCGCTCACCAACT
 ACGACCAAAACACACAATCCTGGAAGTCTCCAACAAAGAAGGAGCTTCCCTCCA
 CCCCTTAACCCATGCTCTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATC
 ATATCAACGCAAACAGCAATACCAGCAACAACCTCTACCTTCCCTGCATCTGTCATC
 ACTTATGAAACCAAACAGAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAG
 CCTCAAAAGAAGAACAGGACTGACCCAAACATCTCCAGACCAGAGTTCTGGCTTC
 ACTACACCCACTGCTATGACACCTCTGCTCTGGCATTCACTCATTCCCCACCAGAA
 AACACAACCTGGATTCAAGCACAATCAGTTTCAAGAACCTTAATCTGACA
 GATGTGATTGAAGAACTAGCCAAAGCAAGTACTCAGACTTTGAAGAGCACAAATTGCT
 TCTGAAACAACTTGTCCAGCAATCACACCAGACTTCAAGAACCTAGGAAAGCATCA
 TTAGACACTCCCATAACCACCAATTCTTGAGCAGCAGTGCTACTCTAACGTTCC
 ATCTCCCTCCCTTACTCAGAGAGCAGTTACTGACACACGTGGCAGTCCATTTC
 CGGCTTATGACAATACAGTGGCAAGCTGCACGAATCCTCAAGGCACAATCTCAA
 ATGCCAAGTTACAATTGGAACCACTCACTCATCTACCTCTAACTGTTACATTCT
 ACTCCCATGCCAGCACTAACAAACAGTTAAATCACAGAATTCCAATTAACCTCCATCT
 CCCTGGGAGAATACCAATTGGCACAAACCAACTCAGACATTGCTGAAAAGGC
 AAAAGCCAGAAGTAAGCATGTTGGCTACTACAGGCCGTGCGAGGCCACCAACTCTT
 GTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGAGTGACTTTGATAAGAAACCA
 GTTCAAGAAGCAACAACTTCCAAACTCCTCCCTTGACTTTGTCTAGGTATATA
 TTTGAAAAGCCAGGATAGTTGGAGGAAAGCTGCAAGTTTACTATTCCAGCTAAC
 TCAGATGCCATTCTCCCTGTGAAGCTGTTGGAAATCCCTGCCACCAATTGTC
 ACCAGAGTTCAAGGACTTGATTATCTAGAGGAAACCAAGAACAGCAGGGTCCAGGTT
 CTCCCAATGGTACCCGTCCATCCAGAGGGTGGAAATTCAAGGACCGCGGACAGTAC
 TTGTGTTCCGCATCCAATCTGTTGGCACAGACCACCTCATGTCACCTTGTCTG
 GTTTCCTATCCTCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTCTATTCC
 GGAAGCACTGTGGAACGTGAAGTGCAAGAGCAGAAGGTAGGCCAAGCCCTACAGTTAC
 TGGATTCTGCAAACAAACAGTTGTCTCAGAATCATCCCAGGGAAAGTAGGCAGGCT
 GTGGTGACGGTTGACGGAACATTGGCTCCACAACTCAGTATTATGACCGTGGC
 TTTTACAAATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTAAAATA
 CAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGC
 ACTTGGGTGAAAGTTAAAACGCCCTGACTGCAAAAGGAACCTCCTCAGCCCAGC

GTTTACTGGGTCTCTGATGGCACTGAAGTGAAACCATTACAGTTACCAATTCC
 AAGTTGTTCTTATTTCAAATGGGACTTGTATATAAGAAACCTAGCCTCTCAGAC
 AGGGGCACCTATGAATGCATTGCTACCAGTCCACTGGTCGGAGCGAAGAGTAGTA
 ATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGATAGAAGCTGCATCCCAGAAA
 AGGACTGAAGTGAATTGGGGACAAATTACTACTGAACTGCTCAGCCACTGGGGAG
 CCCAAACCCCAAATAATGTGGAGGTACCATCCAAGGCTGTGGCGACCAGTGGAGC
 TGGATCCACGTCTACCTAATGGATCCCTGTTATTGGATCAGTAACAGAAAAAGAC
 AGTGGTGTCTACTTGTGTGGCAAGAAACAAATGGGGATGATCTGATACTGATG
 CATGTTAGCCTAAGACTGAAACCTGCCAAATTGACCACAAGCAGTATTTAGAAAG
 CAAGTGCTCCATGGGAAAGATTCCAAGTAGATTGCAAAGCTCCGGCTCCCCAGTG
 CCAGAGATATCTGGAGTTGCCTGATGGAACCATGATCAACAATGCAATGCAAGCC
 GATGACAGTGGCCACAGGACTAGGAGATACCCCTTCAACAATGGAACCTTATAC
 TTCAACAAAGTTGGGTAGCGGAGGAAGGAGATTATACTGCTATGCCAGAACACC
 CTAGGGAAAGATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATA
 AGGCAGAGTAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAGCTGTCCTGAC
 TGTGAGGTCACTGGGATCCAAACCAAAATATTTGGTGTGCTGCCCTCCAATGAC
 ATGATTTCCCTCTCCATTGATAGGTACACATTCTGATGCCAATGGGTCTTGACCATC
 AACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCAGT
 GGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAAT
 GGTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCAAAAAAA
 CACTTTGACTGCAGAGCTGAAGGGACACCATCCTGAAGTCATGTGGATCATGCCA
 GACAATATTTCTCACAGCCCCATACTATGGAAGCAGAACATCACAGTCCATAAAAAT
 GGAACCTGGAAATTAGGAATGTGAGGCTTCAGATTAGCCGACTTTATCTGTGTG
 GCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATG
 CTGAGAAGACCGACATTAGAAATCCATTAAATGAAAAAAATAGTTGCCCAGCTGGGAA
 AAGTCCACAGCATTGAATTGCTCTGTTGATGGTAACCACCCACCTGAAATAATCTGG
 ATTTACCAAATGGCACACGATTTCAATGGACCACAAAGTTATCAGTATCTGATA
 GCAAGCAATGGTTCTTTATCATTCTAAAACAACTCGGGAGGATGCAGGAAAATAT
 CGCTGTGCAGCTAGGAATAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATT
 GCCCAGAAGCCAGTTATTCTACCTATGCACCAGGGACAGTAAAGGCATCAGTGGAA
 GAATCTCTATCACTGCATTGTGTCTGATGGAATCCCTAACGGCAAATATCAAATGG
 ACTATGCCAAGTGGTTATGTAGTAGACAGGCCTCAAATTAAATGGGAAATACATATTG
 CATGACAATGGCACCTAGTCATTAAGAAGCAACAGCTTATGACAGAGGAAACTAT
 ATCTGTAAGGCTAAAATAGTGTGGTCATACACTGATTACTGTTCCAGTAATGATT
 GTAGCCTACCCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGGACA
 GGGGCAGCCTTCAGCTCCACTGTGTGGCTTGGAGTTCCAGGCCAGAAATCACA
 TGGGAGATGCCTGACCAACTCCCTCTCAACGGCAAGTAAAGAGAGGACACATGGAA
 AGTGAGCAGCTTCACTTACAAGGTACCCCTAGTCATTAGAATCCCCAACCTCCGAT
 TCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGTAGTGTGATTATGCAGCAACG
 TATATTCAAGTAATCTGA

Figure 54 Human OCP: predicted amino acid sequence

MKVKGRGITC LLVSFAVICL VATPGGKACP RRCACYMPTE VHCTFRYLT
 IPDSIPPNE RINLGYNSLV RLMETDFSL TKLELLMLHS NGIHTIPDKT
 FSDLQALQVL KMSYNKVRKL QKDTFYGLRS LTRLHMDHNN IEFINPEVFY
 GLNFLRLVHL EGNQLTKLHP DTFVSLSYLQ IFKISFIKFL YLSDNFLTSL
 PQEMVSYMPD LDSLYLHGNP WTCDCHLKWL SDWIQEKPDV IKCKKDRSPS
 SAQQCPLCMN PRTSKGKPLA MVSAAAFQCA KPTIDSSLKS KSLTILEDSS
 SAFISPQGFM APFGSLTLNM TDQSGNEANM VCSIQKPSRT SPIAFTEEND
 YIVLNTSFST FLVCNIDYGH IQPVWQILAL YSDSPLILER SHLLSETPQL
 YYKYKQVAPK PEDIFTNIEA DLRADPSWLM QDQISLQLNR TATTFSTLQI
 QYSSDAQITL PRAEMRPVKH KWTMISRDNN TKLEHTVLVG GTVGLNCPGQ
 GDPTPHVDWL LADGSKVRAP YVSEDGRILY DKSGKLELQM ADSFDTGVYH
 CISSNYDDAD ILTYRITVVE PLVEAYQENG IHHTVFIGET LDLPCHSTGI
 PDASISWVIP GNNVLYQSSR DKVVLNNGL RILQVTPKDQ GYYRCVAANP
 SGVDFLIFQV SVKMKGQRPL EHDGETEGSG LDESNPIAHL KEPPGAQLRT
 SALMEAEVGK HTSSTSKRHN YRELTQLQRRG DSTHRRFREN RRHFPPSARR
 IDPQHWAALL EKAKKNAMPD KRENTTVSPP PVVTQLPNIP GEEDDSSGML
 ALHEEFMVPA TKALNLPART VTADSRTISD SPMTNINYGT EFSPVVNSQI
 LPPEEPTDFK LSTAIKTTAM SKNINPTMSS QIQGTTNQHS STVFPLLLGA
 TEFQDSDQMG RGREHFQSRP PITVRTMIKD VNVKMLSSTT NKLLLESVNT
 TNSHQTSVRE VSEPRHNHY SHTTQILSTS TFPSPDPTHAA HSQFPIPRNS
 TVNIPLFRRF GRQRKIGGRG RIISPYRTPV LRRHRYSIFR STTRGSSEKS
 TTAFSATVLN VTCLSCLPRE RLTTATAALS FPSAAPITFP KADIARVPSE
 ESTTLVQNPL LLLENKPSVE KTTPTIKYFR TEISQVTPTG AVMTYAPTSI
 PMEKTHKVNA SYPRVSSTNE AKRDSVITSS LSGAITKPPM TIIAITRFSR
 RKIPWQQNFV NNHNPKGRLR NQHKVSLQKS TAVMLPKTSP ALPQRQSSPF
 HFTTLSTSVM QIPSNTLTTA HHTTTKTHNP GSLPTKKELP FPPLNPMLPS
 IIISKDSSTKS IIISTQTAIPA TTPTFPASVI TYETQTERSR AQTIQREQEP
 QKKNRTDPNI SPDQSSGFTT PTAMTPPALA FTHSPPENTT GISSTISFHS
 RTLNLTDVIE ELAQASTQTL KSTIASETTL SSKSHQSTTT RKASLDTPIP
 PFLSSSATLM PVPISPPFQ RAVTDTRGDS HFRLMTNTVV KLHESSRHNL
 QMPSSQLEPL TSSTSNLHS TPMALTTVK SQNSKLTPSP WAEYQFWHKP
 YSDIAEKGKK PEVSMMLATTG LSEATTLVSD WDQOKNTKKS DFDKKPVQEA
 TTSKLLPFDS LSRYIFEKPR IVGGKAASFT IPANSDAFLP CEAVGNPLPT
 IHWTRVSGLD LSRGNQNSRV QVLPNGTLSI QRVEIQDRGQ YLCASASNLF
 TDHLHVTLSV VSYPPIRILER RTKEITVHSG STVELKCRAE GRPSPTVTWI
 LANQTVVSES SQGSRQAVVT VDGTQLVLHNL SIYDRGFYKC VASNPQGQDS
 LLVKIQVIAA PPVILEQRRQ VIVGTWGESL KLPCTAKGTP QPSVYWLSD
 GTEVKPLQFT NSKLFLFSNG TLYIRNLASS DRGTYECIAT SSTGSERRVV
 MLTMEERVTS PRIEAASQKR TEVNFGDKLL LNCSATGEPK PQIMWRLPSK
 AVVDQWSWIH VYPNGSLFIG SVTEKDSGVY LCVARNKMGD DLILMHVSLR
 LKPAKIDHKQ YFRKQVLHGK DFQVDCKASG SPVPEISWSL PDGTMINNAM
 QADDSGHRTR RYTLFNNGL YFNKVGVAEE GDYTCYAQNT LGKDEMVKHL
 TVITAAPRIR QSNKTNKRIK AGDTAVLDCE VTGDPKPKIF WLLPSNDMIS
 FSIDRYTFHA NGSLTINKVK LLDSEYVCV ARNPSGDDTK MYKLDVVSXP
 PLINGLYTNR TVIKATAVRH SKKHFDCAE GTPSPEVMWI MPDNIFLTAP
 YYGSRITVHK NGTLEIRNVR LSDSADFICV ARNEGGEESVL VVQLEVLEML
 RRPTFRNPFN EKIVAQLGKS TALNCSVDGN PPPEIIWILP NGTRFSNGPQ
 SYQYLIASNG SFIISKTTRE DAGKYRCAAR NKVGYIEKLV ILEIGQKPV

90/90

LTYAPGTVKG ISGESLSLHC VSDGIPKPNI KWTMPSGYVV DRPQINGKYI
LHDNGTLVIK EATAYDRGNY ICKAQNSVGH TLITVPVMIV AYPPRITNRP
PRSIVTRTGA AFQLHCVALG VPKPEITWEM PDHSLLSTAS KERTHGSEQL
HLQGTLVIQN PQTSDSGIYK CTAKNPLGSD YAATYIQVI*